



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2011

Evaluation de la réforme de la maturité 1995 (EVAMAR). Rapport final de la phase II

Eberle, F ; Gehrer, K ; Jaggi, B ; Kottonau, J ; Oepke, M ; Pflüger, M ; Huber, C ; Husfeldt, V ;
Lehmann, L ; Quesel, C

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-60193>

Published Research Report

Published Version

Originally published at:

Eberle, F; Gehrer, K; Jaggi, B; Kottonau, J; Oepke, M; Pflüger, M; Huber, C; Husfeldt, V; Lehmann, L;
Quesel, C (2011). Evaluation de la réforme de la maturité 1995 (EVAMAR). Rapport final de la phase
II. Bern: Secrétariat d'Etat à l'éducation et à la recherche SER.



Evaluation de la réforme de la maturité 1995 (EVAMAR)

Rapport final de la phase II

Réalisée à la demande de la Conférence suisse des
directeurs cantonaux de l'instruction publique CDIP et
du Secrétariat d'Etat à l'éducation et à la recherche SER



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de l'intérieur DFI
Secrétariat d'Etat à l'éducation et à la recherche SER

Dans sa collection «Dossiers SER», le Secrétariat d'Etat à l'éducation et à la recherche SER publie des études prospectives, des résultats de recherche, des rapports sur des thèmes d'actualité dans les domaines de l'éducation et de la recherche. Mis ainsi à la disposition d'un large public, ces textes participeront au débat sur ces questions. Les textes publiés n'engagent pas nécessairement le Secrétariat d'Etat à l'éducation et à la recherche.

© 2011 Secrétariat d'Etat à l'éducation et à la recherche SER

ISSN: 1662-2634



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de l'intérieur DFI
Secrétariat d'Etat à l'éducation et à la recherche SER
Education

Hallwylstrasse 4
CH - 3003 Berne
T +41 31 322 96 57
F +41 31 322 78 54
info@sbf.admin.ch
www.sbf.admin.ch

Evaluation de la réforme de la maturité 1995 (EVAMAR)

Rapport final de la phase II

Réalisée à la demande de la Conférence suisse des directeurs cantonaux de l'instruction publique CDIP et du Secrétariat d'Etat à l'éducation et à la recherche SER

Franz Eberle, Karin Gehrler, Beat Jaggi,
Johannes Kottonau, Maren Oepke, Michael Pflüger

Institut de pédagogie gymnasiale et professionnelle,
Université de Zurich
(équipe principale du projet)

Christina Huber, Vera Husfeldt, Lukas Lehmann,
Carsten Quesel

Haute école spécialisée du Nord-Ouest de la Suisse,
Haute école pédagogique
(auteurs du chapitre VI)

Table des matières

Avant-propos du directeur du projet	17
L'essentiel en bref	18
I Introduction au projet global	25
1 Réforme de la formation gymnasiale	25
2 EVAMAR I	27
3 Autres études d'évaluation	29
4 Révision partielle du RRM de 1995	30
5 Mandat et problématique d'EVAMAR II	30
5.1 Concept global d'évaluation	30
5.2 Enquêtes retenues dans le mandat définitif pour EVAMAR II	34
5.3 Limites de l'étude EVAMAR II	36
5.4 Organisation de projet et remerciements	38
II Module A: éléments de l'aptitude aux études selon les hautes écoles	40
1 Base de données des sous-projets A1, A3 et A4	41
2 Sous-projets A1 et A3: analyse de contenu des supports d'études et des examens	42
2.1 Méthode et instruments du SP A1 et du SP A3	42
2.1.1 Méthode du codage fin	42
2.1.2 Méthode du codage large	45
2.1.3 Qualité des techniques de mesure	46
2.2 Retour des données dans les sous-projets A1 et A3	48
2.3 Evaluations dans les sous-projets A1 et A3	48
2.4 Résultats de l'évaluation des sous-projets A1 et A3; interprétation	49
3 Sous-projet A4: enquête auprès des enseignants	50
3.1 Le questionnaire	50
3.2 Base de données du SP A4 et déroulement de l'enquête	51
3.3 Résultats de l'enquête du SP A4	51
3.3.1 Signification des disciplines gymnasiales pour les enseignements (question 1)	51

3.3.2	Contenus concrets des connaissances et capacités dans les branches gymnasiales importantes pour l'enseignement visé (question 2)	56
3.3.3	Lacunes actuellement constatées dans les connaissances et capacités (question 3)	59
3.3.4	Importance des compétences interdisciplinaires pour l'enseignement (question 4)	61
3.3.5	Carences constatées dans les connaissances et capacités interdisciplinaires (question 5)	63
3.3.6	Connaissances et capacités de départ supplémentaires souhaitables	63
3.3.7	Conclusions	66
4	Sous-projet A2: analyse des examens d'entrée	67
4.1	Aperçu	67
4.2	Le test d'aptitude aux études de médecine	67
4.3	L'examen d'admission à la HSG	68
4.4	Conclusions	68
III	Module B: Modélisation des connaissances et capacités requises pour entreprendre des études universitaires dans certaines disciplines	69
1	Généralités	69
2	La grille de compétences en langue première	71
2.1	Bases	71
2.2	Grilles et modèles de compétences existants pour la langue première	72
2.2.1	Grille zurichoise d'analyse de textes (Zürcher Textanalyseraster)	72
2.2.1.1	Brève présentation	72
2.2.1.2	Appréciation de la grille zurichoise de textes pour une utilisation dans EVAMAR II	73
2.2.2	DESI	73
2.2.2.1	Brève présentation	73
2.2.2.2	Appréciation du modèle de compétence de DESI pour son emploi éventuel dans EVAMAR II	75
2.2.3	PISA	76
2.2.3.1	Brève présentation	76
2.2.3.2	Appréciation du modèle PISA pour une utilisation éventuelle dans EVAMAR	78
2.2.4	Le Cadre européen commun de référence CECR	79
2.2.4.1	Brève présentation	79
2.2.4.2	Appréciation du CECR pour une utilisation effective dans EVAMAR II	83
2.2.5	HarmoS	84
2.2.5.1	Brève présentation	84

	2.2.5.2	Appréciation du modèle HarmoS pour son utilisation dans EVAMAR II	86
	2.2.6	Relevé de l'état des prestations scolaires dans le land de Rhénanie-du-Nord-Westphalie (NRW)	86
	2.2.6.1	Brève présentation	86
	2.2.6.2	Appréciation du relevé des prestations scolaires en NRW pour son utilisation dans EVAMAR II	87
	2.3	Modèle final de la grille de compétences «langue première» pour EVAMAR II	89
	2.3.1	Motivation du choix du CECR comme base de travail	89
	2.3.2	Présentation de la grille de compétences retenue par EVAMAR pour la langue première	90
	2.4	Synthèse	93
3		La grille de compétences en mathématique	95
	3.1	Bases	95
	3.2	Grilles de modèles de compétences existants pour les mathématiques	96
	3.2.1	Le modèle de compétences du NCTM (National Council of Teachers of Mathematics)	96
	3.2.2	Le modèle de compétences de la KMK en Allemagne	96
	3.2.3	Le modèle HarmoS de compétences en mathématiques	97
	3.2.4	Le modèle PISA de compétences en mathématiques	99
	3.2.5	Modèles TIMSS	99
	3.3	Modèle final de la grille de compétences «mathématiques» pour EVAMAR II	101
4		La grille de compétences en biologie	105
	4.1	Bases	105
	4.2	Grilles et modèles de compétences existants pour la biologie	105
	4.2.1	La classification allemande de Ratisbonne (RVK)	105
	4.2.2	Comparaison de la structure de manuels standard pour l'enseignement gymnasial en allemand	106
	4.2.3	Comparaison de la structure de manuels standards pour les étudiants débutant en biologie	107
	4.2.4	Les EPA Biologie et les standards de formation en Allemagne	110
	4.2.4.1	EPA Biologie	110
	4.2.4.2	Standards de formation KMK en biologie	113
	4.2.5	Résultat intermédiaire pour la dimension des contenus: synthèse	114
	4.2.6	Culture scientifique	118
	4.2.7	Standards de formation KMK	118
	4.2.8	EPA Biologie	119
	4.3	Description de la grille de compétences EVAMAR en biologie	120

IV	Module C: développement et application de tests	122
1	Généralités et vue d'ensemble	122
1.1	Objectifs et choix de l'objet des enquêtes	122
1.2	Observations sur la détermination des domaines de test	123
1.2.1	De la valeur prédictive des tests pour l'aptitude aux études supérieures	123
1.2.2	Choix de la biologie comme branche des sciences naturelles	128
2	Description des instruments de relevé	129
2.1	Démarche suivie pour la construction des tests: remarques générales	129
2.2	Description de l'évaluation qualitative des experts et des prétests (y compris procédure de révision et modification des items)	131
2.3	Les traductions dans les deux autres langues nationales: généralités pour les trois disciplines testées	131
2.4	Démarche suivie pour la détermination des items intégrés dans l'analyse générale	132
2.5	Le test de langue première	133
2.5.1	Précisions sur l'élaboration du test	133
2.5.1.1	A propos de la construction des items	133
2.5.1.2	A propos de la traduction des items	138
2.5.2	Description des cahiers de test	139
2.5.3	Caractérisation des items intégrés dans l'analyse générale et ancrage dans la grille de compétences	139
2.6	Le test de mathématiques	140
2.6.1	Précisions sur l'élaboration du test	140
2.6.1.1	A propos de la construction des items	140
2.6.1.2	A propos de la traduction des items	141
2.6.2	Description des cahiers de test	141
2.6.3	Format des items et intégration dans la grille de compétences	142
2.7	Le test de biologie	142
2.7.1	Précisions sur l'élaboration du test	142
2.7.1.1	A propos de la construction des items	142
2.7.1.2	A propos de la traduction des items	146
2.7.2	Description des cahiers de test	146
2.7.3	Format des items et intégration dans la grille de compétences	146
2.8	Le test des compétences transversales (TCT)	147
2.8.1	Description des bases du test	147
2.8.2	Description des cahiers de test	148

	2.8.3	Format des items	148
2.9		Le questionnaire complémentaire	148
	2.9.1	Raison d'être et contenu	148
	2.9.2	Description du questionnaire	149
	2.9.3	Qualité technique du questionnaire	149
3		Définition de l'échantillon	150
	3.1	Définition de la population cible	150
	3.2	Constitution de l'échantillon	151
	3.3	Taux de retour	151
4		Collecte et traitement des données	152
	4.1	Déroulement de la collecte des données	152
	4.2	Codage des réponses aux tests	152
5		Résultats nationaux	153
	5.1	Généralités	153
	5.2	Résultats de langue première	154
	5.2.1	Résultats d'ensemble	154
	5.2.2	Résultats d'items particuliers	156
	5.3	Résultats de mathématiques	161
	5.3.1	Résultats d'ensemble	161
	5.3.2	Résultats d'items particuliers	163
	5.4	Résultats de biologie	169
	5.4.1	Résultats d'ensemble	169
	5.4.2	Résultats d'items particuliers	170
	5.5	Comparaison avec les notes de maturité	178
	5.6	Résultats du test des compétences transversales (TCT)	179
	5.7	Résultats du questionnaire complémentaire	180
	5.7.1	Perception des tests	180
	5.7.2	Perception des compétences transversales propres	181
	5.7.3	Intérêt pour les différentes disciplines et valeur propédeutique du gymnase	182
	5.8	Interprétation des résultats sous l'angle de l'aptitude aux études	183
6		Analyses comparatives selon plusieurs critères	185
	6.1	Généralités	185
	6.2	Comparaisons selon la strate géographique	186
	6.2.1	Résultats des disciplines testées	186

6.2.1.1	Langue première	186
6.2.1.2	Mathématiques	188
6.2.1.3	Biologie	190
6.2.2	Notes de maturité	193
6.2.3	Résultats du test des compétences transversales (TCT)	194
6.2.4	Résultats choisis issus du questionnaire complémentaire	195
6.3	Comparaisons selon l'option spécifique choisie	196
6.3.1	Résultats des disciplines testées	197
6.3.1.1	Langue première	197
6.3.1.2	Mathématiques	198
6.3.1.3	Biologie	200
6.3.2	Notes de maturité	203
6.3.3	Résultats du test des compétences transversales	206
6.3.4	Résultats choisis issus du questionnaire complémentaire	208
6.4	Comparaisons selon le sexe	210
6.4.1	Résultats de langue première	210
6.4.2	Résultats de mathématiques	211
6.4.3	Résultats de biologie	212
6.4.4	Résultats du test des compétences transversales	212
6.4.5	Résultats choisis issus du questionnaire complémentaire	213
6.5	Comparaisons selon le taux de maturités	213
6.6	Comparaisons selon la durée de la voie gymnasiale	216
6.7	Autres comparaisons	219
6.7.1	Assiduité et résultats des tests	219
6.7.2	Examens oraux et écrits	220
7	Corrélations	221
7.1	Corrélations des tests entre eux	221
7.2	Corrélations entre les tests et les notes de maturité	222
7.3	Corrélations des notes de maturité entre elles	223
8	Synthèse	224
8.1	Bases des tests et de l'enquête complémentaire	224
8.2	Résultats d'ensemble	224
8.3	Résultats des tests du point de vue structurel	225
8.4	Pratiques d'évaluation	226
8.5	Perception des compétences transversales propres	227

V	Sous-projet D1: analyse d'examens écrits de maturité	228
1	Généralités et vue d'ensemble	228
1.1	Objectifs et choix des disciplines	228
1.2	Généralités sur les instruments d'analyse	228
2	Description de l'échantillon	232
3	Résultats	232
3.1	Langue première	232
3.1.1	Les outils d'analyse	232
3.1.2	Les contenus des examens en langue première	233
3.1.2.1	La forme de l'épreuve	233
3.1.2.2	Les textes de départ soumis aux élèves	236
3.1.2.3	Datation des textes	239
3.1.2.4	Auteurs (hommes et femmes)	240
3.1.2.5	Thématique	242
3.1.2.6	Longueur des textes soumis	243
3.1.2.7	Possibilités de choix	244
3.1.2.8	Explications complémentaires	245
3.1.2.9	Moyens auxiliaires	245
3.1.2.10	Durée de l'examen	246
3.1.3	Exigences cognitives	246
3.1.4	Comparaison avec les tâches du module C	249
3.1.5	Comparaison avec les résultats du sous-projet A4	250
3.1.6	Appréciation globale des examens de maturité en langue première en Suisse	252
3.2	Mathématiques	253
3.2.1	Les instruments d'analyse	253
3.2.2	Les contenus des épreuves d'examen en mathématiques	254
3.2.2.1	Démarche	254
3.2.2.2	Résultats	255
3.2.3	Les exigences cognitives	256
3.2.4	Comparaison avec les items du module C	262
3.2.5	Comparaison avec les résultats du sous-projet A4	262
3.2.6	Appréciation globale des examens de maturité en mathématiques en Suisse	266
3.3	Biologie	266
3.3.1	Les instruments d'analyse	267

3.3.2	Les contenus des épreuves d'examen en biologie	268
3.3.3	Exigences cognitives	269
3.3.4	Comparaison avec les tâches du module C	277
3.3.5	Comparaison avec les résultats du sous-projet A4	279
3.3.6	Evaluation globale des examens de maturité en biologie en Suisse	280
VI	Sous-projet D2: la qualité des travaux de maturité en Suisse	282
1	Contexte général et bases théoriques	283
1.1	Le travail de maturité dans la réforme de 1995	283
1.1.1	Introduction du travail de maturité	283
1.1.2	Interdisciplinarité du travail de maturité	283
1.1.3	Le travail de maturité dans l'examen de maturité	284
1.2	Le travail de maturité dans la formation gymnasiale	284
1.2.1	Aptitude aux études supérieures et maturité intellectuelle	286
1.2.2	Propédeutique scientifique	287
1.2.3	Compétences transversales	288
1.2.4	Bilan partiel	289
1.3	Etat de la recherche sur le travail de maturité	289
1.3.1	Typologie des travaux de maturité	290
1.3.2	Choix du sujet	290
1.3.3	Réalisation	291
1.3.4	Savoirs, compétences et savoir-faire des élèves	291
1.3.5	Suivi du travail de maturité	291
1.3.6	Bilan partiel	292
2	Bases méthodologiques et réalisation de l'enquête	292
2.1	Schéma de production d'un travail de maturité	292
2.2	Evaluation des travaux de maturité	293
2.2.1	Dimensionnement de l'enquête	293
2.2.2	Normes de référence servant à juger de la qualité des travaux de maturité	294
2.2.3	Conception de l'instrument d'analyse	296
2.2.4	Notation	297
2.2.5	Contrôle anti-plagiat	297
2.3	Comparaison des conditions de réalisation des travaux	298
2.4	Enquête auprès des élèves	298
2.5	Vérification de l'hypothèse de fonctionnement	299

2.5.1	Qualité du travail de maturité	299
2.5.2	Niveau institutionnel	300
2.5.3	Niveau du suivi	301
2.5.4	Niveau individuel	302
2.6	Population étudiée et échantillonnage	303
2.6.1	Population étudiée	303
2.6.2	Echantillonnage	304
2.6.3	Collecte des données	306
3	Cadres institutionnels	306
3.1	Directives générales	306
3.1.1	Directives nationales	306
3.1.2	Objectifs cantonaux	307
3.2	Directives émanant des établissements	309
3.2.1	Attributions et directives	309
3.2.2	Objectifs définis par les établissements	309
3.2.3	Préparation et choix du sujet	311
3.2.4	Consignes formelles	312
3.2.5	Mode opératoire et calendrier	313
3.2.6	Suivi du travail de maturité	314
3.2.7	Evaluation et notation	315
3.3	Bilan partiel	316
4	Le travail de maturité vu par les élèves	318
4.1	Choix du sujet	318
4.2	Rattachement du travail de maturité à une matière et choix du sujet	319
4.3	Choix du sujet du travail de maturité	320
4.4	Evaluation du travail de maturité	322
4.5	Suivi des travaux	323
4.6	Cadre institutionnel et situation personnelle	330
4.7	Appréciation du volume de travail de préparation et de l'utilité du travail de maturité	332
4.8	Bilan partiel	338
5	La valeur de propédeutique scientifique du travail de maturité	338
5.1	Double évaluation	338
5.2	Caractéristiques de l'échantillon	339
5.3	Conception des échelles	341

5.4	Qualité du fond	344
5.5	Qualité de la langue	345
5.6	Qualité de la forme	346
5.7	Appréciation globale	347
5.8	Aspects individuels et institutionnels de la qualité	347
5.9	Bilan partiel	350
6	Conclusions	352
	Références bibliographiques du chapitre VI	354
	Tableaux et figures du chapitre VI	361
	Annexe du chapitre VI	363
	Annexe VI.A: Vue d'ensemble des hypothèses testées	363
	Annexe VI.B: Grille d'évaluation des conditions cadre du travail de maturité	365
	Annexe VI.C: Facteurs de pondération	366
VII	Résumé	367
1	Objets de l'étude	367
2	Résultats I: Identification des performances à mesurer (pré-analyse)	368
3	Modélisation des compétences (B)	372
4	Résultats II: Mesure des performances (C)	372
5	Résultats III: Analyse des examens de maturité	379
6	Résultats IV: Analyse des travaux de maturité	380
7	Les limites de l'étude EVAMAR II	382
VIII	Conclusions et recommandations du chef de projet	383
1	Des résultats généraux satisfaisants	383
2	Assurer l'aptitude générale aux études supérieures	384
3	Maintenir un large éventail de disciplines et le droit général à l'admission aux universités	387
4	Assurer une compensation à l'entrée dans les filières d'études	388

5	Développer la capacité d'apprendre de manière autonome sans restreindre l'éventail des disciplines	388
6	Améliorer la maîtrise de la langue première	389
7	Faire davantage référence à des standards dans les programmes d'enseignement	389
8	Faire davantage référence à des standards dans l'évaluation des performances	390
9	Maintenir le gymnase en quatre ans	390
10	Assurer l'acquisition d'une aptitude aux études supérieures dans toutes les options spécifiques	391
11	Réduire les disparités entre les taux de maturité	392
12	Poursuivre les efforts dans le domaine de l'égalité des sexes	392
13	Développer le travail de maturité	392
14	Poursuivre les analyses	392
	Bibliographie (sans le chapitre VI)	393
	Liste des tableaux et figures	405
	Liste des abréviations	415

Avant-propos du directeur du projet

En été 2001, la Confédération et la CDIP ont décidé de réaliser une évaluation à l'échelle nationale, baptisée EVAMAR, de la réforme de la formation gymnasiale introduite par le RRM en 1995. La première phase (EVAMAR I), fondée sur des questionnaires, a étudié l'adéquation entre l'éventail des disciplines à option et les centres d'intérêt des élèves, la perception subjective de la qualité de la préparation à des études universitaires, la mise en œuvre des objectifs pédagogiques transversaux ainsi que la gestion des réformes par les établissements scolaires. Les résultats de l'évaluation ont donné satisfaction. La deuxième phase, EVAMAR II, a commencé à la fin 2005 et s'est conclue à l'automne 2008; l'analyse a porté sur d'autres éléments de la formation gymnasiale. Elle a mis l'accent sur un relevé objectif du niveau de formation des élèves à la fin du gymnase face aux exigences des études universitaires. Même s'il n'est guère possible de mesurer précisément ou de prévoir pour l'ensemble des bacheliers et bacheliers, l'aptitude à réussir des hautes études quelles qu'elles puissent être, les constats du présent rapport jettent des éclairages utiles sur divers aspects et résultats de la formation gymnasiale dans les écoles de notre pays.

Mes remerciements vont d'abord à tous les bacheliers, à toutes les bacheliers qui se sont mis à disposition pour effectuer des tests et répondre aux questionnaires ou qui ont mis à disposition leurs travaux de maturité pour l'évaluation. Je remercie aussi tout particulièrement toutes les personnes - enseignants, membres de directions d'école et collaborateurs administratifs - qui nous ont apporté un appui actif dans le relevé des données.

Notre reconnaissance va aussi aux nombreuses personnes qui ont pris part à notre projet et qui sont nommées à la fin du premier chapitre. Sans leur grand labeur, il eût été impossible de mener à bien un projet d'une aussi grande complexité.

Le rapport que vous avez entre les mains a eu plusieurs rédacteurs et rédactrices, qui ont aussi mené les recherches et analyses dont ils rapportent les résultats. Cette organisation explique que les chapitres et sections du rapport présentent des différences d'ordre stylistique ou, selon les auteurs ou les sujets, structurelles et visuelles.

Le chapitre VI est intégralement l'œuvre des collaborateurs de la Haute école spécialisée du Nord-Ouest de la Suisse, Haute école pédagogique. Ils ont conçu et conduit entièrement le sous-projet en question.

Zurich, le 31 octobre 2008

Franz Eberle

L'essentiel en bref

En été 2001, la Confédération et la Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP) ont décidé de réaliser une évaluation à l'échelle nationale, baptisée EVAMAR, de la réforme de la formation gymnasiale introduite par le RRM en 1995. La première phase de cette évaluation (EVAMAR I) a porté essentiellement sur les trois thèmes suivants: (1) l'adéquation entre l'éventail de disciplines à option offert et les centres d'intérêt des élèves et son incidence sur le succès de la formation, notamment en termes de perception subjective de la qualité de la préparation à des études universitaires; (2) la mise en œuvre des objectifs pédagogiques transversaux; (3) la gestion de la réforme par les écoles. Cette phase a eu pour instrument privilégié le questionnaire et n'a pas inclus de mesures de performances. Les résultats d'EVAMAR I, publiés fin 2004, montraient que la réforme dans son ensemble donnait satisfaction. En été 2005, la Confédération et la CDIP ont décidé de passer à la seconde phase de l'évaluation, EVAMAR II, dont l'élément principal devait être un relevé objectif du niveau de compétences des élèves à la fin du gymnase.

Question de recherche

Un relevé évaluatif du niveau de formation doit logiquement se référer aux objectifs de cette formation. Selon l'art. 5 du RRM 95, la formation gymnasiale doit premièrement permettre d'acquérir la «maturité requise pour entreprendre des études supérieures» (maturité nécessaire ou aptitude aux études supérieures) et, deuxièmement, préparer les élèves à «assumer des responsabilités au sein de la société actuelle» (maturité sociale étoffée d'une vaste culture générale). Les incidences de ces deux objectifs sur les programmes sont parfois importantes et vont tantôt dans le même sens, tantôt dans des sens divergents. Il serait extrêmement coûteux d'opérationnaliser entièrement ces objectifs en compétences concrètes que les gymnasiennes et les gymnasiens sont censés acquérir dans toutes les disciplines et tous les domaines interdisciplinaires et de mesurer systématiquement le niveau réel de chacun, comme cela était prévu dans le premier projet proposé par le mandataire. C'est la raison pour laquelle le mandant a choisi de limiter les enquêtes en concrétisant certains éléments seulement de l'aptitude aux études universitaires, qui est l'objectif principal de la formation gymnasiale. Le second objectif (maturité sociale étoffée d'une vaste culture générale) n'a de ce fait été inclus qu'en partie (aspects communs seulement) dans la définition opérationnelle des objectifs choisie par le mandant.

Bases et pré-analyses

Voici les bases sur lesquelles se fondaient les tests et questionnaires soumis à un échantillon d'environ 3800 bacheliers et bacheliers de Suisse alémanique et de Suisse romande, ainsi que l'analyse d'examens écrits de maturité provenant des trois régions linguistiques: analyse des supports didactiques et des premiers examens intermédiaires dans les 16 disciplines d'études universitaires les plus fréquentées de Suisse, enquête auprès des professeurs enseignant ces branches (cours de bases), prise en compte des théories et recherches les plus récentes en matière de mesure des performances cognitives et de prédiction de la réussite des études.

Le test de langue première met l'accent principalement sur les compétences linguistiques transversales qui jouent un rôle essentiel pour toutes les voies d'études.

Le test de mathématiques fait intervenir avant tout des contenus utiles pour un grand nombre de branches et non uniquement pour les études de mathématiques.

Le test de biologie est un test de connaissance sur des contenus utiles surtout en biologie ou en médecine, mais dans une certaine mesure aussi pour les branches ne faisant pas partie des sciences expérimentales (par ex. la psychologie).

Le test de compétences transversales (TCT) se réfère à l'examen d'aptitude aux études de médecine et mesure des aptitudes cognitives assez générales. Il n'a été réalisé qu'à titre complémentaire en Suisse alémanique.

La batterie de tests comprend donc différents types d'indicateurs prédictifs de la réussite des études et peut de ce fait être qualifiée d'équilibrée. Si l'on comprend l'aptitude générale aux études universitaires comme la possibilité d'entreprendre sans problème des études dans n'importe quelle branche, on accordera aux compétences mesurées dans les domaines testés de notre enquête l'importance suivante: la maîtrise de la langue première fait partie des compétences de base de chaque branche d'études ou presque; les mathématiques sont importantes pour un grand nombre de branches faisant appel au langage mathématique formel; les connaissances de biologie ne sont requises que pour un nombre restreint de branches, mais elles sont des bases essentielles pour les études susmentionnées; et les compétences transversales mesurées avec le TCT sont en corrélation certaine avec la réussite des études de médecine, ce qui vaut aussi pour toutes les autres orientations exigeant des aptitudes identiques ou similaires. Les pré-analyses d'EVAMAR II ont en outre démontré l'avantage qu'il y avait à posséder, même à un niveau différent, un certain bagage dans la discipline de maturité correspondant à la branche d'études choisie.

Résultats généraux des tests

En moyenne, environ 50% des items ont été résolus correctement; les résultats étaient légèrement meilleurs en langue première qu'en mathématiques et en biologie. Dans cette dernière discipline, les résultats obtenus par le groupe de bacheliers et bacheliers qui n'avaient plus de cours à leur programme depuis une demi-année ou plus étaient clairement moins bons, sans doute sous l'effet de l'oubli. Cela démontre que les connaissances acquises dans les écoles de maturité ne restent pleinement présentes que peu de temps et que l'idéal est de conserver une discipline gymnasiale au programme le plus longtemps possible pour entreprendre ses études supérieures rapidement. Les résultats du TCT n'étaient que légèrement inférieurs à ceux de l'examen d'aptitude aux études de médecine effectué chaque année, ce qui est tout à fait plausible étant donné que l'enquête EVAMAR a pris en compte également toutes les personnes qui ne pensaient pas avoir les aptitudes nécessaires pour faire des études de médecine. Ce résultat est, de plus, un indicateur important de la grande motivation dont les bacheliers et bacheliers ont fait preuve lors de la réalisation du test EVAMAR. La proximité des résultats du TCT par rapport à ceux atteints dans le cadre de l'examen d'aptitude, qui est autrement plus déterminant pour le parcours de l'individu, montre que ce dernier ne le cède en rien aux situations d'évaluation notée habituelles au gymnase.

Les résultats diffèrent cependant beaucoup, tant entre individus qu'entre classes entières. Les différences constatées sont plus importantes que celles observées dans les notes effectivement délivrées à la maturité. On peut en déduire qu'une proportion considérable de bacheliers et bacheliers possèdent des compétences probablement insuffisantes dans au moins l'une des matières examinées lors de notre test. On parvient au même constat en analysant les notes de maturité décernées par les écoles elles-mêmes, surtout en mathématiques. 41,4% des bacheliers et bacheliers ont en effet obtenu une note insuffisante à l'écrit de mathématiques en 2007. Ils étaient encore 24,4% à la note finale de la maturité, qui tient compte également de la notation continue et de l'examen oral. On doit donc partir du principe que tous les bacheliers et bacheliers ne possèdent pas de compétences suffisantes sur l'ensemble de la gamme pour entreprendre n'importe quelles études sans avoir encore à pallier des lacunes, alors même que l'obtention du certificat de maturité a fait d'eux des personnes «aptes aux études de n'importe quel type». Ce constat concorde avec les compétences lacunaires en mathématiques et en langue première observées subjectivement par les professeurs. Le système de compensation des notes insuffisantes lors des examens de maturité rend toutefois quasi inévitable cet effet de la formation gymnasiale.

Analyse structurelle des résultats des tests

Le nombre d'années de formation gymnasiale passées dans un établissement de ce type semble avoir un impact sur les résultats des tests. Une période de quatre ans au moins dans un gymnase va de pair avec des résultats meilleurs que ceux obtenus dans le cas de structures où la première des quatre années s'effectue dans une classe «gymnasiale» de l'école secondaire locale. Ce phénomène est clair pour les mathématiques et légèrement perceptible également pour la langue première et, en Suisse alémanique, pour le TCT.

Les gymnases «longue durée» se classent mieux que les autres dans presque tous les domaines de test, légèrement seulement il est vrai. C'est en effet uniquement en biologie que l'on observe des différences allant de moyennes à assez élevées en faveur des gymnases «longue durée», même si les notations continues y sont proportionnellement moins bonnes, légèrement certes, mais tout de même de manière significative. Il se pourrait cependant qu'il s'agisse là plutôt d'une anomalie ou du résultat d'une notation en fonction de la classe.

On constate des différences parfois considérables entre les divers groupes d'option spécifique. Il fallait s'attendre, en raison du système, à ce que les meilleurs résultats en mathématiques soient obtenus par le groupe de l'option spécifique Physique et applications des mathématiques (PAM) et en biologie par le groupe de l'option spécifique Biologie et chimie (BCH). Mais l'excellent résultat au test de langue première obtenu par les bacheliers et bachelères de l'option spécifique Langues anciennes par comparaison à celui du groupe Langues modernes, qui n'a atteint qu'un rang moyen, surprend quelque peu. Il n'y a pas non plus d'explication d'ordre systémique pour les autres variations observées dans quelques domaines de compétence. C'est le groupe de l'option spécifique Langues anciennes qui s'est le mieux classé en termes de régularité, tandis que les bacheliers et bachelères des trois nouvelles options spécifiques Musique, Arts visuels et Philosophie/pédagogie/psychologie (PPP) se sont montrés plutôt mauvais. L'analyse comparative des notes de maturité effectivement décernées et des résultats du test fait apparaître quelques différences de classement et surtout des écarts moins marqués. Il est frappant de voir que certains groupes d'option spécifique avaient à peu près 50% (Langues modernes, Arts visuels, Musique), voire plus de 50% (PPP) de notes insuffisantes à l'écrit de mathématiques. Pour la note finale de maturité en mathématiques, ces groupes d'option (Musique exceptée cette fois-ci) avaient encore environ 30% de notes insuffisantes. Le meilleur groupe au TCT était de loin celui de l'option PAM. Les groupes des nouvelles options PPP et Arts visuels de même que, pour une matière, Musique étaient quant à eux en queue de classement.

Les différences constatées entre les sexes rejoignent en grande partie les résultats des études précédentes. Les bachelères se sont plutôt mieux classées en langue première, les bacheliers clairement mieux en mathématiques et dans la partie du TCT consacrée aux sciences expérimentales.

Notation

Les notes attribuées ne coïncident qu'en partie avec les résultats du test. Une première explication à cela est que ce dernier mesurait des compétences partiellement différentes de celles vérifiées par les examens de maturité. Ces différences sont dues à la cible plus restreinte de l'enquête et n'ont rien à voir avec une réalisation lacunaire des objectifs du gymnase. Une deuxième explication de ce constat se situe dans la procédure de notation. On peut d'une part envisager qu'il y ait une adaptation partielle de la notation au potentiel de la classe (référence à la norme sociale), ce qui, par rapport à une notation se référant à des standards de qualité interclasses (référence aux critères), mène à des notations différentes. Mais on peut d'autre part, en comparant les notes, constater également certaines convergences structurelles avec les résultats du test, ce qui donne à penser que les enseignantes et les enseignants se réfèrent pour une part considérable dans leur notation à des «standards de qualité objectivés». On trouve par ailleurs dans la notation certaines anomalies peu explicables, qui laissent elles aussi supposer l'existence d'une pratique faisant seulement partiellement référence au critère. Citons par exemple la notation continue en biologie, plus haute dans les groupes d'option spécifique Langues anciennes et Arts visuels que dans le groupe Biologie et chimie. Il est enfin étonnant que les examens écrits de maturité, en

mathématiques surtout, mais aussi en langue première, donnent lieu à une forte proportion de résultats insuffisants, mais soient chaque fois «remontés» grâce à la note de l'oral et à la notation continue.

Examens écrits de maturité

Dans toutes les disciplines faisant l'objet de l'enquête, les examens écrits de maturité sont pour une part souvent exigeants, ils requièrent pour leur réalisation la mise en œuvre de diverses aptitudes cognitives (il ne suffit pas de simplement restituer des connaissances préalablement apprises par cœur) et couvrent des domaines importants de l'aptitude aux études. Mais nous avons trouvé d'autre part bon nombre d'examens ne correspondant pas à ce profil. Dans l'ensemble, les énoncés sont donc assez hétérogènes.

Travaux de maturité

La qualité propédeutique des travaux de maturité examinés peut être majoritairement qualifiée de satisfaisante. Bien que sa durabilité n'ait pas encore été prouvée pour l'instant, bon nombre d'indices suggèrent que le travail de maturité est une forme d'apprentissage et d'examen utile et profitable en vue de l'aptitude aux études supérieures.

Les limites de l'étude EVAMAR II

Il y a lieu de rappeler ici expressément les limites de l'étude EVAMAR II. Il faudra en tenir compte quand il s'agira d'interpréter les résultats et d'en déduire des mesures concernant l'organisation future du gymnase. En principe, le projet ne permet de commenter que certains aspects choisis des objectifs de la formation gymnasiale. Voici les éléments relevant d'une aptitude générale aux études supérieures qu'EVAMAR II n'a pas étudiés et dont il n'a pas analysé la présence chez les bacheliers et bacheliers: premièrement les connaissances et compétences qui ne sont pas directement exigées à l'université, bien qu'elles fassent partie d'une vaste culture générale, permettant notamment d'assumer ses responsabilités au sein de la société (maturité sociale), et qui doivent donc s'acquérir dans le cadre de la formation gymnasiale; deuxièmement les connaissances et compétences qui font nécessairement partie de l'aptitude aux études supérieures, mais qui sont issues d'autres disciplines que la langue première, les mathématiques et la biologie; et troisièmement les éléments de compétences transversales que le cadre défini pour le projet n'aurait permis de saisir qu'imparfaitement, malgré leur importance pour la réussite des études (par ex. la capacité d'apprendre de manière autonome, de s'organiser de manière générale, de faire des recherches, d'assimiler rapidement une riche bibliographie spécialisée, etc.). La portée de ces facteurs ne pourrait être établie qu'à l'aide d'autres enquêtes, proposées dans le concept général de l'étude (études longitudinales auprès des étudiants) et il faudrait pour la mesurer des procédures d'évaluation lourdes à réaliser. Rappelons toutefois que ces facteurs ont déjà été étudiés et validés approximativement dans le cadre du projet EVAMAR I par le biais de questionnaires, autrement dit par autoévaluation. Le questionnaire joint à EVAMAR II comprenait en outre des questions à ce sujet, complémentaires à EVAMAR I.

Conclusions et recommandations principales du directeur du projet

1. Les bacheliers et bacheliers de Suisse ont obtenu dans les domaines analysés des résultats qui sont normaux, s'agissant de tests développés de façon standard pour une population déterminée. Au concept de normalité se rattache le fait que le taux de réussite aux items proposés a été en moyenne de 50%. Dans ces conditions, on peut considérer le niveau de compétences comme satisfaisant. Cela ne revient pourtant pas à affirmer que tous les bacheliers et bacheliers disposent de compétences suffisantes dans toute la gamme des disciplines. Cette réserve confirme une hypothèse déjà faite naguère au vu de la proportion relativement importante de notes insuffisantes mises aux examens de maturité, principalement dans le domaine des mathématiques et de l'épreuve écrite de langue première. Donc, malgré le bon niveau de

compétences, des améliorations sont possibles en principe. Une telle conclusion n'a rien de surprenant, car elle ressort de toutes les mesures de performances comportant des tests standardisés – à quelque degré scolaire que ce soit. Il serait urgent d'agir si les tâches proposées avaient donné lieu à une proportion nettement inférieure à 50% de réponses correctes et que l'ensemble des aptitudes individuelles s'était situé nettement au-dessous de la moyenne. Cela n'est pas le cas chez les bacheliers et bacheliers de notre pays.

2. Si l'on veut interpréter la notion d'aptitude générale aux études universitaires en postulant que le certificat de maturité doit sanctionner des compétences suffisantes à l'entrée pour entreprendre des études dans n'importe quelle branche, alors on doit préciser que cela n'est pas le cas pour une bonne partie des bacheliers et bacheliers. Ce constat d'EVAMAR II concerne les seuls domaines analysés, mais le tableau ne serait probablement guère différent avec d'autres disciplines. Ici, nous n'avons rien observé de nouveau non plus, et les lacunes dont nous parlons se manifestent aussi dans les notes de maturité délivrées en fonction de la possibilité de compensation prévue par le RRM. Par là même, on explique aussi les points de vue critiques des chargés d'enseignement universitaire – nombreux à déplorer des lacunes chez leurs étudiants, principalement en langue première et en mathématiques – , ainsi que les taux d'échec parfois élevés lors des premiers examens d'université. La question fondamentale est de savoir s'il est possible de viser à l'avenir une aptitude universitaire sans restriction pour tous les bacheliers et bacheliers. Un tel objectif exigerait une révision des critères de réussite aux examens de maturité, où l'on interdirait les notes de maturité insuffisantes. S'ensuivraient des effets indésirables, notamment une baisse du taux de maturités, et une adaptation partielle de la mise des notes à la situation nouvelle. Une «maturité universitaire sans faille» pour tous les bacheliers et bacheliers est donc du domaine de l'utopie. Nous sommes cependant bien loin de recommander l'abandon de l'exigence posée au certificat de maturité de sanctionner une aptitude universitaire générale et de donner un accès sans examen aux hautes écoles. La motivation détaillée sur ce point figure dans le corps du rapport. Notre position se fonde en particulier sur le fait que le système en vigueur n'autorise pas n'importe quel profil de notes insuffisantes dans les domaines de compétences. Les bacheliers et bacheliers ayant des performances insuffisantes dans certaines branches doivent bel et bien les compenser doublement par de bons résultats dans les autres disciplines. En outre, les effets réels des compétences lacunaires doivent être pris en compte: EVAMAR I et d'autres recherches ont montré que le choix des études s'opère sur la base des aptitudes personnelles et des intérêts qui y sont liés, et qu'ainsi les lacunes constatées dans certaines disciplines ne se répercutent que d'une manière limitée sur la réussite des études entreprises. Sans doute y aurait-il lieu, dans la perspective d'une révision du RRM, d'envisager de soumettre les compétences de base importantes pour de nombreuses voies d'études à des critères de réussite plus sévères, par ex. à l'impossibilité générale de compenser les résultats insuffisants. Une telle disposition toucherait par exemple certaines compétences en langue première et en mathématiques: non pas la totalité des savoir-faire du programme de ces disciplines, mais celles qui concernent essentiellement la capacité aux études. On songe ici à certains secteurs de la langue première et des mathématiques, mais non d'autres contenus tout aussi importants dans le cursus gymnasial. (Il ne s'agit donc pas d'exiger d'une manière générale une note minimale de 4 en mathématiques). Une telle modification redonnerait du crédit à l'idée que la certification aboutissant à la maturité reconnaît globalement l'aptitude aux études universitaires quelles qu'elles soient. En outre, on peut supposer que des critères de réussite stricts sans possibilité de compensation auraient un effet favorable sur le pilotage des études des études universitaires: les bacheliers et bacheliers se résigneraient moins vite à «faire leurs calculs» en comptant avec une insuffisance dans certains domaines. Comme l'anglais, en sa qualité de langue de communication scientifique au niveau universitaire, appartient aussi au domaine des compétences essentielles à l'aptitude aux études supérieures, – et cet avis est partagé par les universitaires – il conviendrait de rendre obligatoire pour l'anglais aussi l'accession à une compétence standard encore à définir.
3. Les enquêtes auprès des professeurs, coordonnées avec les questions complémentaires posées aux bacheliers et bacheliers, confirment les analyses menées jusqu'ici, dans la mesure où les intéressés eux-

mêmes, en situation d'autoévaluation, qualifient souvent de lacunaire leur degré d'autonomie dans l'apprentissage et le travail. Il convient donc d'appuyer les recommandations de ces études: les gymnases devraient accorder une importance encore accrue à la promotion ciblée de l'autonomie dans l'apprentissage. On favoriserait ainsi une compétence générale essentielle pour la poursuite des études, et on aiderait les bacheliers et bacheliers à combler eux-mêmes des lacunes individuelles inévitables, selon les disciplines choisies, perçues au début de leurs études. Quant aux déficits constatés dans les compétences de langue première, ils pourraient être diminués en insistant de manière encore plus systématique sur l'emploi correct de la langue standard dans toutes les disciplines.

4. Un encouragement accru de l'apprentissage et du travail individuels autonomes ne doit pas se faire aux dépens de l'étendue des contenus spécifiques transmis dans l'enseignement des disciplines, ni en limitant l'éventail des branches offertes dans les gymnases. On risquerait alors de voir apparaître de nouvelles lacunes restreignant le choix des études, du fait que le large éventail universitaire ne serait plus réellement ouvert à tous; l'autre objectif de la maturité selon le RRM (maturité sociale par une culture générale approfondie, qui n'est pas l'objet des analyses d'EVAMAR II) pourrait aussi s'en trouver compromis. Cette difficulté postule que l'on fasse l'effort, dans toutes les disciplines, de se concentrer encore mieux sur l'essentiel, notamment en procédant à une analyse affinée de ce qui est indispensable au niveau des prérequis d'une maturité sociale approfondie et en améliorant la concertation des gymnases avec les hautes écoles sur ce que ces dernières attendent des nouveaux étudiants.
5. Par ailleurs, pour remédier aux difficultés liées à l'impossible «maturité universitaire sans faille», on peut souhaiter l'augmentation de l'offre des cours d'appui que les universités mettent sur pied dans telle ou telle discipline selon les filières. Des exemples existent déjà.
6. On a constaté des différences parfois considérables d'une classe de gymnase à l'autre. Les notes confirment généralement la tendance des résultats obtenus aux tests. On a noté pourtant des «anomalies». S'il est clair que les enseignants des gymnases savent se référer à des standards de qualité transdisciplinaires, on n'en doit pas moins supposer qu'ils s'adaptent au niveau de performances effectif de leurs classes respectives. Ce phénomène obère la valeur prédictive des notes décernées à la maturité. Il serait donc judicieux d'examiner les moyens de garantir une meilleure référence aux standards tout en déconseillant plutôt les examens centralisés, qui comportent, à n'en pas douter, beaucoup d'aspects problématiques.
7. Les résultats indiquent que les établissements connaissant un cursus gymnasial continu sur trois ans sont défavorisés. Il conviendrait de repenser la disposition concernée du RRM. Les données d'EVAMAR II révèlent aussi que les écoles comportant une structure de «sous-gymnase» – donc une durée de gymnase encore allongée – obtiennent en fin de parcours un niveau supérieur pour toutes les prestations relevées (notes de maturité comprises); l'écart est léger, mais significatif cependant. Un tel fait devrait être pris en compte dans les discussions, déjà engagées dans plusieurs endroits, sur une éventuelle suppression du gymnase de longue durée.
8. Les différences de performances relevées entre les divers groupes d'options spécifiques sont souvent considérables. Elles apparaissent non seulement dans les tests passés par EVAMAR II, mais aussi, quoique dans une proportion moindre, dans les notes de maturité délivrées par les écoles elles-mêmes. Il serait indiqué, d'une part, de rechercher les causes de cet état de faits par le biais d'analyses longitudinales, quand bien même les conséquences concrètes ne sont pas d'une extrême gravité pour les bacheliers et bacheliers, du fait du choix «adapté» de leur voie d'études. D'autre part, il conviendrait aussi d'étudier des mesures propres à améliorer les résultats de la formation dans des domaines partiels, essentiels pour les études, des disciplines «mathématiques» et de la «langue première». Ce souci concerne surtout les élèves ayant suivi les «filières nouvelles» des OS introduites avec le RRM – et non représentées dans les types de maturité qui existaient auparavant.

9. L'hypothèse de travail qu'à un taux de maturité élevé correspondent des résultats inférieurs s'est trouvée renforcée non seulement dans les tests, mais aussi souvent dans les notes. L'inégalité manifeste des chances est un problème qui doit faire l'objet d'une discussion politique.
10. Les différences selon les sexes sont conformes à une image déjà connue. Il s'agit de poursuivre les efforts déployés pour améliorer les performances des gymnasiennes en mathématiques et en sciences expérimentales, et d'envisager des enquêtes ultérieures qui porteraient sur d'autres différences éventuelles dans des disciplines qu'EVAMAR II n'a pas analysées.
11. La forme d'apprentissage que constitue le travail de maturité a fait ses preuves; l'élaboration d'un projet personnel d'une certaine durée est un élément central de l'apprentissage gymnasial et qui mérite d'être développé encore - dans une mesure modeste au vu des facteurs de réussite recensés dans les analyses d'EVAMAR II.
12. Le projet EVAMAR II se présentait comme un projet transversal: aussi ses résultats posent-ils plus de nouvelles questions qu'ils n'apportent de réponses sûres sur les causalités. Les questions en suspens mériteront d'être étudiées de façon approfondie, grâce à des analyses longitudinales qui utiliseront systématiquement les données personnelles recueillies au cours de la formation gymnasiale et des études ultérieures.

I Introduction au projet global

1 Réforme de la formation gymnasiale

La formation gymnasiale suisse, fondamentalement réformée en 1995, a reçu une nouvelle réglementation dans le RRM, le règlement de reconnaissance de la maturité (CDIP, 1995). La maturité suisse donne l'accès général illimité à toutes les études de niveau universitaire. Un numerus clausus n'existe que pour les études de médecine, dont l'accès implique un test d'aptitude en complément à la qualification de la maturité (Hänsen & Spicher, 2007). La réforme de la formation de maturité devait l'adapter aux nouvelles exigences tout en maintenant la garantie d'accès à toutes les filières d'études des universités et des hautes écoles polytechniques.

Jusqu'à la réforme de 1995, les études gymnasiales suisses permettaient de choisir, entre cinq types de maturité (A à E), celui qui correspondait aux aptitudes et à l'inclination. L'idée centrale de la réforme de 1995 était de remplacer ces types de maturité par la «maturité unique», assortie d'un large éventail de choix et de possibilités d'individualisation.

Les principales caractéristiques de la maturité unique sont les suivantes:

- La maturité unique comporte un nombre relativement élevé de disciplines fondamentales, égales quant à l'essentiel pour tous les gymnasiens, afin de garantir une formation équilibrée et largement diversifiée. Il s'agit des disciplines suivantes: la langue première, une deuxième langue nationale, une troisième langue (soit une troisième langue nationale, soit l'anglais, soit une langue ancienne), les mathématiques, les sciences expérimentales (comportant un enseignement obligatoire de la biologie, de la chimie et de la physique), les sciences humaines et sociales (comprenant un enseignement obligatoire de l'histoire, de la géographie et une introduction à l'économie et au droit), les arts visuels et/ou la musique.
- Tous les élèves doivent en outre choisir une option spécifique et une option complémentaire et ils sont tenus de réaliser un travail de maturité.

Les caractéristiques principales de l'individualisation sont les suivantes:

- L'option spécifique peut être choisie entre huit variantes: langues anciennes (latin et/ou grec), une langue moderne (une troisième langue nationale, l'anglais, l'espagnol ou le russe), physique et applications des mathématiques, biologie et chimie, économie et droit, philosophie/pédagogie/psychologie, arts visuels, musique.
- L'option complémentaire peut être choisie entre 13 variantes: physique, chimie, biologie, applications des mathématiques, histoire, géographie, philosophie, enseignement religieux, économie et droit, pédagogie/psychologie, arts visuels, musique, sport.
- Le thème du travail de maturité peut être déterminé individuellement, d'entente avec l'enseignante ou l'enseignant en charge de l'encadrement; le travail de maturité peut être pluridisciplinaire.

Parmi les autres directives du RRM (CDIP, 1995), citons les suivants:

- Le certificat de maturité ne peut être décerné qu'au terme de 12 ans d'études, dont quatre ans dans une école de maturité. Exceptionnellement, trois ans de gymnase suffisent, lorsque le degré secondaire I comporte un «enseignement de caractère prégymnasial» (cette réglementation est controversée dans le milieu des gymnases).
- Les examens de maturité sont réussis à condition que le double de la somme de tous les écarts vers le bas par rapport à la note 4 ne soit pas supérieur à la somme simple de tous les écarts vers le haut par rapport à cette

même note. En d'autres termes, pour pouvoir compenser une note insuffisante de 3, il faudra obtenir la note 5 dans deux autres disciplines.

Les gymnases suisses ne sont soumis ni à des normes contraignantes en matière d'enseignement ou de formation ni à un contrôle centralisé de la réalisation des objectifs. On ne trouve des directives matérielles que dans l'article directeur du RRM (CDIP, 1995) et dans le plan d'études cadre (PEC) (CDIP, 1994) de la Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP).

Les contenus essentiels de l'objectif des études sont résumés dans la citation suivante (CDIP, 1995, art. 5): «Ces écoles dispensent une formation générale équilibrée et cohérente, qui confère aux élèves la maturité requise pour entreprendre des études supérieures et les prépare à assumer des responsabilités au sein de la société. Elles évitent la spécialisation et l'anticipation de connaissances et d'aptitudes professionnelles (...)». Ce passage formule les deux objectifs d'études prioritaires du gymnase, à savoir le développement de la maturité requise pour entreprendre des études supérieures et la préparation à la prise en charge de responsabilité au sein de la société. La deuxième phrase indique en outre fondamentalement la voie à suivre pour atteindre ces objectifs (cf. Eberle, 2008, quant à des réflexions sur l'adéquation, la pertinence scientifique et la cohérence de l'article sur l'objectif des études).

On trouve dans le PEC, sous une forme un peu plus différenciée, mais encore peu fonctionnalisée, tant la description des objectifs interdisciplinaires que celle des objectifs de disciplines individuelles. Les objectifs de la première catégorie, dits «objectifs de compétences interdisciplinaires», qui sont remarquables, comprennent les composantes suivantes (CDIP, 1994, 12):

- compétences sociales, éthiques et politiques;
- compétences intellectuelles, scientifiques et épistémologiques;
- compétences communicatives, culturelles et esthétiques;
- compétences concernant le développement de la personnalité et la santé;
- compétences concernant les méthodes de travail, l'accès aux savoirs et les technologies de l'information.

Dans ce contexte, on distingue les compétences fondamentales pour l'ensemble des jeunes des compétences particulières, spécifiques aux gymnasiens.

La présentation des objectifs par discipline s'articule, pour chacune d'elles, en «objectifs généraux», «considérations, explications» et «objectifs fondamentaux». Ces objectifs fondamentaux sont quant à eux subdivisés en «connaissances», «savoir-faire» et «attitudes». Il n'est pas possible, en l'occurrence, d'analyser en détail les objectifs et les faiblesses du PEC (en particulier le manque de fondement théorique des objectifs de compétences interdisciplinaires, de même que l'assise et la cohérence insuffisantes des objectifs spécifiques aux disciplines).

En ce qui concerne les principes décrits, il ne s'agit que de directives cadre; quant aux périodes d'enseignement attribuées aux différentes disciplines, elles ne sont qu'esquissées par les fourchettes de proportions du RRM 95. Le plan d'études à l'échelle nationale est expressément désigné comme un plan d'études cadre, chacun des 26 cantons et demi-cantons ayant entrepris sa propre mise en œuvre du PEC. Dans les grands cantons, comme celui de Zurich, les divers établissements gymnasiens disposent même de leurs propres plans d'études. Cette variété de programmes se reflète aussi dans les examens de maturité: des examens uniformes ne sont organisés ni sur le plan fédéral, ni sur le plan cantonal. Il est tout à fait courant que des personnes qui enseignent parallèlement la même discipline dans le même établissement concoctent chacune son propre examen de maturité et qu'elles fixent elles-mêmes l'échelle d'évaluation. Malgré la diversité potentielle des types de qualification ou de développement qui en résulte pour les gymnasiens suisses, la maturité qui prévaut en Suisse est une telle maturité «locale», qui donne accès à toutes les universités du pays.

Ainsi, après la réforme de la maturité de 1995 également, les gymnases suisses ne recevront explicitement des services centralisés que des objectifs vaguement contraignants. Ce point ne signifie toutefois pas que des normes décentralisées ou locales n'existent pas, fût-ce au moins implicitement (cf. Oelkers, 2006):

- Le personnel enseignant a généralement une conception professionnelle de son métier, axée sur des normes spécifiques aux disciplines.
- Les moyens d'enseignement tendent à exercer un effet normatif; cependant, les moyens d'enseignement ne sont pas uniformes.
- Les taux d'admission explicites ou implicites dans les gymnases fixent des normes quant aux conditions à remplir par les élèves.

Manifestement, toutefois, il subsiste des différences entre les cantons et entre les établissements scolaires. Citons seulement les suivants:

- Le nombre d'heures affectées aux diverses disciplines diffère considérablement dans certains cas.
- La fourchette des taux de diplôme de maturité dans les cantons (proportion de jeunes gens qui obtiennent la maturité gymnasiale) était comprise, pour la volée de 2007, entre 11,2% dans le canton d'Obwald et 29,5% dans le canton du Tessin, la moyenne nationale suisse étant de 19,2% (Office fédéral de la statistique, 2008).

Le certificat de maturité devait renseigner sur les qualifications obtenues au gymnase et, partant, également sur la mesure dans laquelle l'objectif de la maturité – l'aptitude aux études supérieures – était atteint. Ces dernières années, l'institution de destination, l'université, a de plus en plus mis en cause cette valeur de pronostic, ce qui s'est traduit par le recours à des procédures propres d'évaluation en médecine, la discipline soumise au *numerus clausus*, et par l'introduction de procédures de sélection sévères dans le cadre d'années d'évaluation («assessment»). Les lacunes mentionnées y contribuent probablement. Les corrélations considérables établies en Allemagne entre les notes de maturité et les notes d'examens universitaires, qui atteignent $\rho = .53$ (cf. la méta-analyse dans Trapman et al., 2007), ne sauraient être transposées sans autres à la Suisse, car les notes utilisées en Allemagne sont plus standardisées. Malheureusement, des données récentes représentatives pour la Suisse font défaut. L'étude qui sera prochainement publiée par l'EPFZ («Maturanoten und Studienerfolg») constituera une exception.

La réforme de 1995 devait être mise en œuvre dans les cantons jusqu'à l'été 2001. Elle a été immédiatement suivie de son évaluation. Outre nombre d'études cantonales, le projet EVAMAR (Evaluation de la réforme de la maturité 1995) a été lancé par les autorités politiques compétentes à l'échelle nationale suisse (Confédération et CDIP). La première phase (EVAMAR I) était achevée à la fin de 2004. En juin 2005, la CDIP et la Confédération ont décidé d'engager une deuxième phase d'évaluation (EVAMAR II), qui couvre la période allant jusqu'à la fin de 2008.

2 EVAMAR I

Pour l'essentiel, la phase I d'EVAMAR a permis d'examiner les trois thèmes suivants (Ramseier et al., 2005): (1) l'adéquation de l'offre de disciplines à option avec les intérêts des élèves et l'incidence sur le succès de la formation, en particulier la qualité perçue subjectivement de la préparation aux études dans les hautes écoles; 2) la mise en œuvre des objectifs pédagogiques interdisciplinaires; 3) la gestion des réformes par les établissements scolaires. Le principal instrument de l'évaluation a été l'enquête conduite auprès de quelque 10 000 candidats et candidates à la maturité, du personnel enseignant et de membres des directions d'écoles. On n'a pas recouru aux mesures des performances. Les résultats d'EVAMAR I sont disponibles depuis la fin de 2004. Voici les principaux résultats des deux premières analyses (Ramseier et al., 2005).

- (1) Résultats de l'auto-évaluation des candidats à la maturité quant à l'adéquation de l'offre de disciplines à option avec les intérêts des élèves et à l'incidence sur le succès de la formation, en particulier la qualité perçue subjectivement de la préparation aux études dans les hautes écoles.
- On observe une individualisation accrue de la formation gymnasiale et une adaptation aux intérêts des élèves; les langues anciennes ont perdu en importance.

- Les possibilités de choix et la couverture des intérêts pour les disciplines augmentent la motivation à apprendre.
- L'auto-évaluation générale des élèves quant à leur préparation aux hautes écoles est bonne. A leur avis, l'objectif de la maturité – les rendre généralement aptes aux études dans les hautes écoles – est atteint.
- Le lien entre le profil d'enseignement (option spécifique et option complémentaire choisies) et les études est assez lâche; la relation la plus étroite peut s'observer en économie, mathématiques/informatique et dans les disciplines «dures» de l'ingénierie.
- Le taux d'échecs aux examens de maturité a augmenté en Suisse alémanique. Parmi les différentes disciplines, les mathématiques constituent l'obstacle majeur.
- Les gymnasiennes sont plus nombreuses que les gymnasiens à réussir, mais les gymnasiens sont plus nombreux que les gymnasiennes à prévoir des études dans les hautes écoles.

(2) Résultats des évaluations de la mise en œuvre des objectifs pédagogiques interdisciplinaires.

- S'agissant des travaux de maturité, les travaux empiriques dominent en Suisse alémanique, tandis que les travaux théoriques sont prépondérants dans le reste de la Suisse. Le personnel enseignant critique le manque de critères d'évaluation suffisants et le manque de temps. Les élèves considèrent que la qualité de l'encadrement et les consignes reçues constituent des points problématiques.
- L'enseignement interdisciplinaire est principalement conduit dans «l'optique transdisciplinaire», à partir de la propre discipline sur les autres, c'est-à-dire généralement sans structure temporelle spécifique. Il reçoit une appréciation globale positive. Seules les mathématiques et les sciences expérimentales suscitent des réserves (le manque de temps étant invoqué en l'occurrence comme principale difficulté).
- La promotion des compétences interdisciplinaires est différemment évaluée quant à son succès: généralement, le personnel enseignant est plus positif que les élèves.

Globalement, la réforme est en majorité jugée satisfaisante.

Dans la perspective de l'intérêt supérieur que représente un système d'enseignement de la maturité correctement fondé et conséquemment mis en œuvre, nous indiquons dans la présente section une question laissée ouverte selon nous dans les analyses de la phase I. Le rapport de projet cite les éléments suivants comme innovations décisives du RRM 95 (1995) par rapport à la réglementation préalable de l'ORM 68 (ordonnance sur la reconnaissance de certificats de maturité, 1968): le remplacement du système des types par un système composé de disciplines fondamentales, à option spécifique et à option complémentaire; la forte pondération de l'enseignement interdisciplinaire et la promotion des compétences interdisciplinaires par l'introduction du travail de maturité et par le regroupement des disciplines scientifiques expérimentales, respectivement des sciences humaines et sociales en deux disciplines fondamentales distinctes. Cette sélection ignore au moins une autre innovation essentielle: les disciplines des sciences humaines et sociales n'ont pas été seulement réunies, elles ont été enrichies d'une discipline d'enseignement supplémentaire: l'économie et le droit. Il faudrait absolument intégrer aussi dans l'évaluation l'introduction d'une nouvelle discipline et de ses mesures corolaires (horaire, nouveaux curricula, formation du personnel enseignant, etc.), même si l'on ne parlait, selon une interprétation restrictive, que de l'extension des contenus de la formation gymnasiale dans la discipline fondamentale des sciences humaines et sociales. Dans les faits, les matières individuelles sont restées, si bien qu'il s'agit d'une nouvelle discipline gymnasiale. Par ailleurs, il conviendrait aussi d'évaluer la mise en œuvre et l'incidence de la nouvelle approche visant à introduire l'informatique dans la formation gymnasiale non plus comme une branche en soi, mais en l'intégrant dans les autres disciplines.

3 Autres études d'évaluation

Outre les nombreuses études d'évaluation cantonales, nous voulons présenter les principaux résultats de deux études également conduites à l'échelle nationale suisse, celles de Notter & Arnold (2003 et 2006). Ces chercheurs ont étudié les mêmes questions qu'EVAMAR I, mais leurs enquêtes ont chacune été conduites, en janvier 2002 et en janvier 2005, auprès d'une cohorte d'étudiants de troisième trimestre dans des universités suisses et non pas parmi des gymnasiens et des gymnasiennes. La première cohorte avait encore obtenu le certificat de maturité en majorité sous l'ancien règlement, tandis que la deuxième avait reçu le diplôme de maturité selon la nouvelle réglementation. Les principaux résultats, comparatifs, de la deuxième étude sont les suivants (Notter & Arnold, 2006, p. 7 ss):

- Les étudiants estiment bonnes à plutôt très bonnes les connaissances et capacités qu'ils ont acquises au degré secondaire dans les domaines de la langue maternelle (grammaire, orthographe, littérature, etc.). Les diplômés de toutes les disciplines fondamentales et de toutes les régions linguistiques estiment que leurs connaissances et capacités en ce domaine sont les plus élevées par rapport aux autres disciplines, et ils les jugent au moins «assez importantes» pour leurs études dans de nombreuses disciplines de formation.
- Les étudiants des deux groupes (obtention de la maturité selon l'ancien, respectivement selon le nouveau règlement) portent un jugement très semblable sur leurs connaissances dans la plupart des disciplines. Les diplômés selon le nouveau règlement de maturité estiment leurs connaissances nettement plus élevées que les étudiants diplômés sous l'ancien règlement dans les seules applications informatiques, et ils les jugent plus basses que ces derniers seulement en langues anciennes (et dans une moindre mesure en physique).
- Les étudiants diplômés selon le nouveau règlement de maturité estiment généralement que leurs compétences interdisciplinaires sont plus élevées que celles des diplômés selon l'ancien règlement de maturité.
- Les étudiants diplômés selon le nouveau règlement de maturité considèrent aussi généralement leurs compétences d'apprentissage avec différentes méthodes d'enseignement et d'apprentissage plus élevées que ne les jugent les diplômés selon l'ancien règlement de maturité. Cette remarque vaut en particulier s'agissant d'apprendre par des travaux autonomes sur le long terme (travail de maturité).
- A l'instar des titulaires du diplôme de maturité selon l'ancien règlement, les étudiants qui ont obtenu leur maturité sous la nouvelle réglementation sont en moyenne satisfaits à très satisfaits de «leur» collège en ce qui concerne la préparation à leurs études. S'agissant de leur développement personnel, les étudiants diplômés selon le nouveau règlement sont un peu plus satisfaits que ceux de l'autre groupe, ce qui s'explique sans doute par la liberté de choix plus large dont ils ont bénéficié.
- Généralement, l'adéquation entre les connaissances et les capacités acquises au collège et les exigences posées par les études semble bonne. Au niveau des connaissances propres aux disciplines, cette observation vaut surtout pour les options spécifiques typiques du domaine d'études visé. Chez les élèves ayant choisi des options spécifiques atypiques du domaine d'études considéré, on constate par contre souvent un écart entre l'évaluation des capacités propres au moment de l'obtention de la maturité et l'importance des disciplines déterminantes du domaine d'études en question [celle-ci étant sous-estimée].
- La comparaison entre l'estimation des capacités propres au moment de l'obtention de la maturité et de l'importance de ces capacités pour la première année d'études fait également apparaître certains écarts. Par exemple, les connaissances d'utilisateur informatique et, dans certains cas, les connaissances de programmation informatique revêtent une certaine importance dans la plupart des domaines d'études, mais les étudiants considèrent que leurs connaissances à cet égard étaient encore nettement inférieures au moment de l'obtention de la maturité, ceci bien qu'ils estiment leurs connaissances en la matière supérieures à celles des diplômés selon l'ancien règlement de maturité.
- Seules les connaissances d'anglais sont jugées avoir une grande importance dans tous les domaines d'études, contrairement aux connaissances des autres langues étrangères.

Globalement, les auteurs parviennent à la conclusion que les écoles du degré secondaire supérieur semblent bien réussir, également après la réforme du règlement de maturité, à conjuguer une large culture générale et la transmission de connaissances spécifiques aux différentes disciplines préparant aux études universitaires dans une branche déterminée.

4 Révision partielle du RRM de 1995

A la demande pressante des organisations gymnasiales, plus qu'en raison des résultats des diverses études d'évaluation conduites, une mini-réforme du RRM de 1995 a été décidée pour 2008. Elle comprend les éléments essentiels suivants (CDIP, 2007).

- Les branches des sciences expérimentales (biologie, physique et chimie) ne seront plus soumises à une évaluation groupée (une seule note) et on procédera à une évaluation distincte dans chacune des disciplines (trois notes).
- La part de l'enseignement pour les disciplines des mathématiques et des sciences expérimentales est portée à 25-35% (actuellement 20-30%).
- L'informatique est désormais intégrée dans le catalogue des options complémentaires.
- L'histoire et la géographie, branches réunies dans le domaine des sciences humaines et sociales, sont mentionnées séparément et font chacune l'objet d'une note propre. L'introduction au droit et à l'économie, qui faisait également partie de ce groupe de disciplines à ce stade, ne recevra certes pas de note, mais elle restera obligatoire.
- Désormais, les cantons ont la possibilité de prévoir la philosophie dans la liste des disciplines fondamentales.
- Le travail de maturité est valorisé, en ce sens que la note qui lui est attribuée fera désormais partie de l'évaluation des examens de maturité.

5 Mandat et problématique d'EVAMAR II

La Confédération et la CDIP ont décidé, à l'été 2005, de lancer la deuxième phase d'évaluation. Complémentaire d'EVAMAR I et d'études similaires, le projet EVAMAR II visait à analyser des aspects supplémentaires de la formation de maturité. Il était prioritaire de recenser objectivement l'état de formation des élèves à la fin du gymnase et d'inclure des mesures de performances. Les enquêtes prévues dans le cadre d'EVAMAR II ont été choisies à partir d'un concept d'évaluation global, que les collaborateurs du projet avaient développé dès l'été 2004 sur mandat de la Confédération et de la CDIP. Le mandat de direction du projet EVAMAR II fut ensuite également confié à l'Institut de pédagogie gymnasiale et professionnelle IGB (Institut für Gymnasial- und Berufspädagogik, anciennement Höheres Lehramt Mittelschulen) de l'Université de Zurich.

Pour faire mieux comprendre EVAMAR II, nous décrivons tout d'abord ci-après les grandes lignes du concept global d'évaluation de la formation de maturité, qui est mentionné ci-dessus (section 5.1). La section 5.2 lui fait suite et présente une description succincte des enquêtes choisies pour EVAMAR II.

5.1 Concept global d'évaluation

L'évaluation idéale d'une réforme consisterait à saisir dans une comparaison avant-après, visant à réaliser une analyse causale, les effets de la réforme dans son ensemble et ceux des diverses mesures prises dans son cadre, également selon les modalités voulues par les cantons, au niveau des écoles et des élèves. Cette démarche n'était pas possible en l'occurrence. Les auteurs du rapport concernant EVAMAR I ont déjà noté que la grande variété

des mesures de réforme cantonales – qu'il s'agisse de la mise en œuvre concrète des directives nationales suisses ou de l'exécution de mesures de réforme supplémentaires – empêchent une comparaison avant-après et que l'objectif d'évaluation doit être limité en conséquence (Ramseier et al., 2005, 38 ss). Nous souscrivons à cet avis. En outre, une comparaison avant-après ne pourrait être réalisée dans une mesure prépondérante qu'à posteriori. Sous l'angle de la méthode de recherche, une telle comparaison serait problématique – d'une part en raison du manque de données ex-ante standardisées et validées, qui font largement défaut, d'autre part en raison de l'influence exercée par des variables incontrôlables pendant la phase de transition (par ex. Bortz & Döring, 2002, 518 ss) – ou à tout le moins extrêmement coûteuse. De ce fait, nous il nous a donc fallu retenir d'emblée une approche transversale, si bien que l'enquête a plutôt le caractère d'un recensement de l'état actuel, dont les causes ne sont pas uniquement dues à la réforme. Dans la perspective des mesures visant au développement futur de la formation gymnasiale, cette conception de l'enquête recèle cependant un potentiel de constats considérable.

Sur la base d'un concept global d'évaluation de la formation de maturité, on a discuté durant la phase préalable à l'attribution du mandat de projet EVAMAR II, sous l'angle des méthodes de recherche et du point de vue pragmatique, un large éventail de diverses enquêtes liées l'une à l'autre. Les diverses options ont été comparées l'une par rapport à l'autre, en tenant compte des ressources limitées et en fixant des priorités différentes. Les éléments essentiels du concept global sont décrits ci-après.

Pour être judicieuse, la saisie de l'état de la formation en vue de son évaluation doit être orientée en fonction des objectifs de cette formation. Les objectifs des écoles de maturité sont d'une part prescrits, du moins en termes généraux, dans les bases légales de la formation gymnasiale et, d'autre part, discutés dans la littérature (par ex. Huber, 1997 et 1998). Les principaux objectifs de la formation gymnasiale sont, selon l'art. 5 du RRM de 1995, de conférer d'une part «la maturité requise pour entreprendre des études supérieures» (maturité nécessaire ou aptitude aux études supérieures) et de préparer d'autre part «à assumer des responsabilités au sein de la société actuelle» (profonde maturité sociale grâce à une vaste culture générale).¹ Dans le mandat de projet et la discussion concernant le gymnase, on trouve parallèlement à l'expression de «maturité requise pour entreprendre des études supérieures» celle «d'aptitude aux études supérieures». Même si une discussion approfondie permettait de relever des différences d'acception dans ces deux expressions, outre des points communs, «maturité requise pour entreprendre des études» et «aptitude aux études supérieures», sont interprétées dans le cadre du projet EVAMAR II comme étant synonymes, au sens de l'ensemble des compétences (connaissances, aptitudes, savoir-faire et dispositions) indispensable pour maîtriser fructueusement, c'est-à-dire pour commencer, poursuivre et achever avec succès des études au niveau des hautes écoles, (cf. Huber, en voie d'impression). L'aptitude générale aux études ou la maturité requise pour étudier signifie alors que lesdites compétences s'étendent, conformément au droit général d'accès aux hautes écoles, à tout cursus d'études dans une haute école et à toutes les disciplines. Le lien entre ce premier objectif partiel de l'aptitude aux études et le deuxième objectif partiel de la préparation à assumer des responsabilités au sein de la société n'est pas clair. Les interprétations varient entre une hiérarchisation de ces objectifs (la réalisation du premier objectif de la maturité requise pour mener à bien les études est supérieure et concourt aussi à la réalisation du second objectif, cf. par ex. Szidat, 2003) et l'équivalence d'objectifs relativement indépendants l'un de l'autre (par ex. Ernst, 2004). Cette deuxième interprétation, qui place l'un et l'autre objectif sur un pied d'égalité, nous paraît plus cohérente. En effet, les deux objectifs sont à certains égards alignés quant à leurs effets, mais ils comportent également des effets autonomes sur l'organisation des cursus, qui peuvent être considérables dans certains cas. Les autres descriptions d'objectifs prévus à l'art. 5 du RRM de 1995 peuvent s'insérer logiquement comme des composantes des deux principaux objectifs cités. Par exemple, la familiarisation «avec la méthodologie scientifique» contribue à l'aptitude aux études supérieures ou la capacité «d'acquérir un savoir nouveau» est d'une grande importance (également au sens de la formation permanente) tant quant à l'aptitude aux études que pour assumer des responsabilités (cf. pour des détails Eberle, 2008).

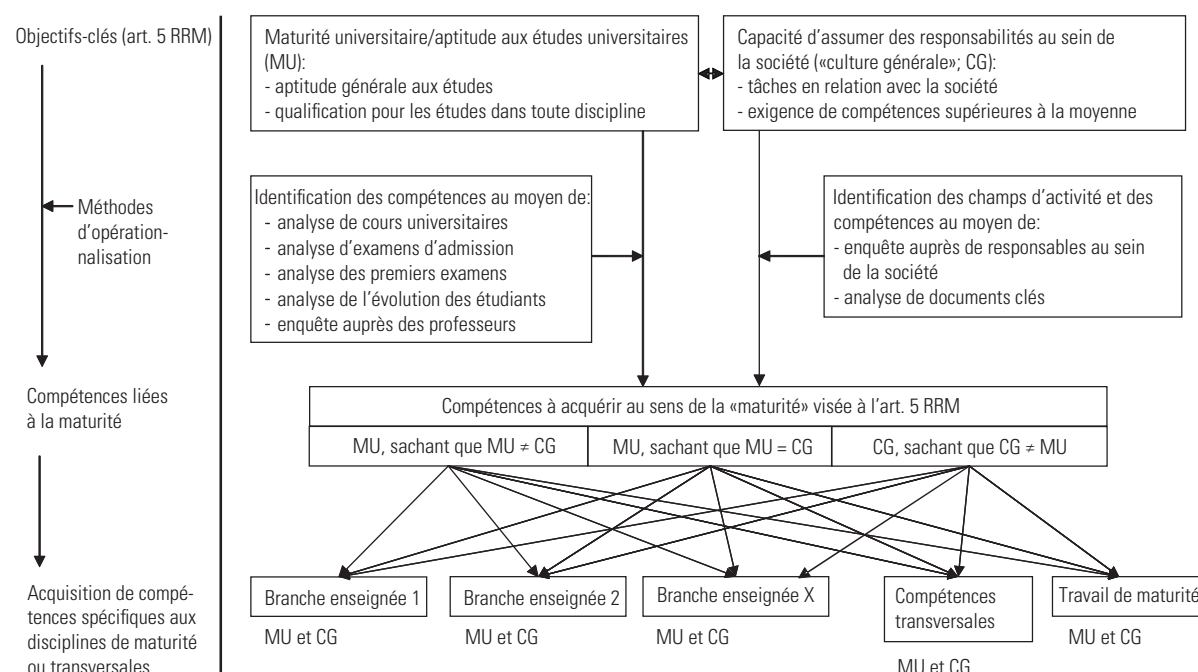
¹ Cette deuxième partie n'est toutefois pas toujours explicitement citée dans les discussions sur la question (par ex. Oelkers, 2008).

Une évaluation complète et axée sur des critères solides doit donc commencer par une modélisation suffisante des objectifs, qui se chevauchent, de l'aptitude aux études et d'une large culture générale. En Suisse, on ne trouve aucune étude dans laquelle ce travail ait été déjà fourni. En Allemagne, il n'existe également pas de telle recherche qui soit fondée sur une base empirique solide (Huber, 1998, 154 ss et Huber, en voie d'impression). Le projet EVAMAR II représentait donc une occasion d'incitation à une telle contribution, qui était attendue depuis longtemps. Les considérations relatives à la concrétisation des objectifs de maturité sont présentées dans la figure I.1. Les méthodes de recherche citées constituent une sélection non exhaustive des méthodes permettant actuellement de déterminer les exigences effectives. Ces méthodes comprennent les composantes suivantes.

- La clarification de l'expression «maturité requise pour entreprendre des études supérieures» ou «aptitude aux études supérieures» du point de vue des hautes écoles par:
 - l'analyse des contenus des cursus, de la documentation, des manuels principaux, etc. dans les universités; la détermination des prérequis spécifiques aux diverses disciplines d'études;
 - l'analyse des contenus des examens d'admission; la détermination des exigences cognitives;
 - l'analyse des contenus des premiers examens dans les universités (par ex. examens intermédiaires; examens du niveau «assessment»; examens du premier semestre dans les EPF); la détermination des exigences cognitives;
 - la détermination des compétences générales d'apprentissage requises par l'analyse des processus d'apprentissage et de pensée des étudiants; des études longitudinale (études de cas): directives de tenue d'un journal d'apprentissage, suivi d'une analyse des contenus; entretiens oraux, « réflexion à haute voix»;
 - les enquêtes auprès de titulaires de chaires, enseignants de cours de base et conseillers aux études des différentes facultés quant aux exigences et aux conditions des études.
- La transposition opérationnelle des «responsabilités au sein de la société» et la déduction des compétences requises pour les assumer, en tant que base de la profonde maturité sociale que permet d'atteindre une vaste culture générale, par:
 - l'enquête et les entretiens d'approfondissement avec des responsables dans la société;
 - l'analyse des contenus de documents comportant des directives normatives pour le développement de la société (par ex. l'Agenda 21 pour un développement durable); l'analyse de modèles de compétences existants comportant des profils d'exigences pour les dirigeants de grandes entreprises; l'analyse de relevés statistiques concernant la carrière professionnelle ultérieure des bacheliers et des bachelières.

A partir des résultats de ces recherches, on pourrait définir concrètement les compétences adaptées aux objectifs que les élèves doivent acquérir dans chaque discipline de maturité et dans chaque domaine interdisciplinaire. Les critères utiles à une évaluation complète de la formation de maturité seraient ainsi déterminés avec suffisamment de précision. Il n'est pas possible d'aborder en l'occurrence les problèmes méthodologiques et normatifs de telles études.

Figure I.1: Modèle global pour la déduction des critères d'évaluation



Légende:

- CG: abréviation pour «culture générale»
- MU: abréviation pour «maturité universitaire»

Clarifier les compétences cibles à acquérir, respectivement déterminer les critères d'évaluation dans la première étape d'évaluation décrite constitue la base de leur utilisation dans la deuxième phase du projet d'évaluation. Des indicateurs les plus divers permettent d'examiner si les valeurs cibles sont atteintes. L'énumération qui suit n'est pas exhaustive:

- résultats de la formation de maturité: développement et réalisation de tests dans toutes les disciplines avec des élèves de gymnase en vue de contrôler les compétences cibles opérationnalisées;
- qualité des plans d'études et des grilles horaires: analyse des contenus des plans d'études et des grilles horaires et comparaison avec les compétences cibles;
- qualité de l'enseignement: évaluation de l'enseignement par l'analyse des contenus des documents didactiques et par des entretiens oraux; éventuellement analyses vidéo;
- qualité des examens de maturité:
 - analyse des contenus des examens de maturité écrits (en particulier sous l'angle des contenus spécifiques aux branches et du niveau d'exigence cognitive) et comparaison des résultats avec les compétences cibles à acquérir dans les diverses disciplines et avec les résultats des tests réalisés (cf. ci-dessus); analyses vidéo d'examens oraux;
 - analyse des contenus des travaux de maturité concernant la répartition des thèmes et des disciplines, le processus de travail des étudiants, la forme et la qualité du travail, le lien aux techniques de travail scientifiques fondamentales comme l'étude des sources et leur traitement correct, l'évaluation du travail, la comparabilité, entre autres, des enquêtes sur le rôle et la fonction du personnel d'encadrement;
 - comparaison des notes d'examens de maturité avec les résultats des examens dans les universités (analyses statistiques).

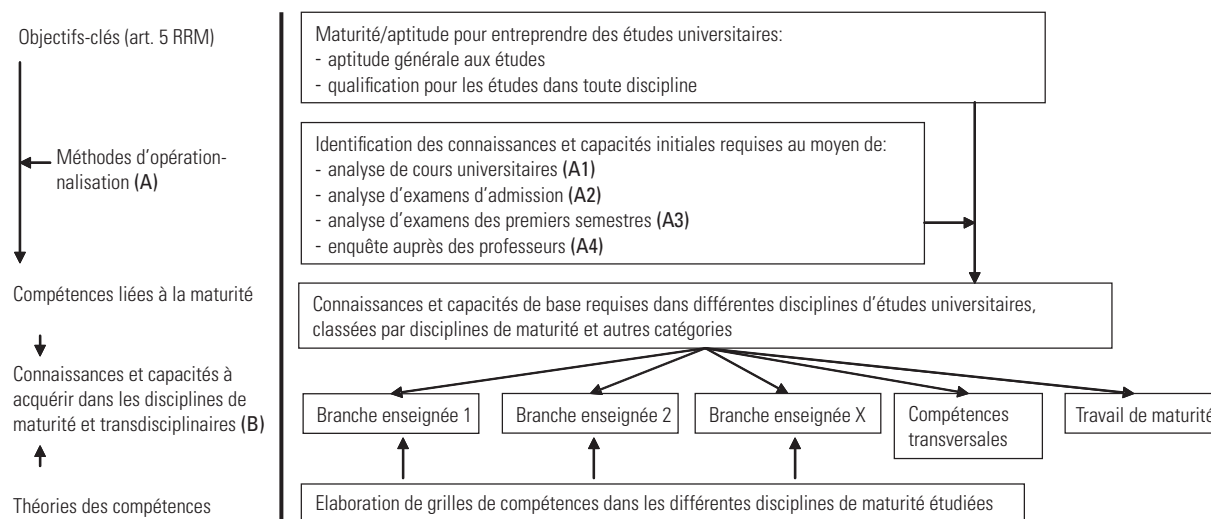
Ces analyses devraient explicitement porter, notamment, sur les innovations que constituent l'introduction à l'économie et au droit et l'informatique intégrée.

5.2 Enquêtes retenues dans le mandat définitif pour EVAMAR II

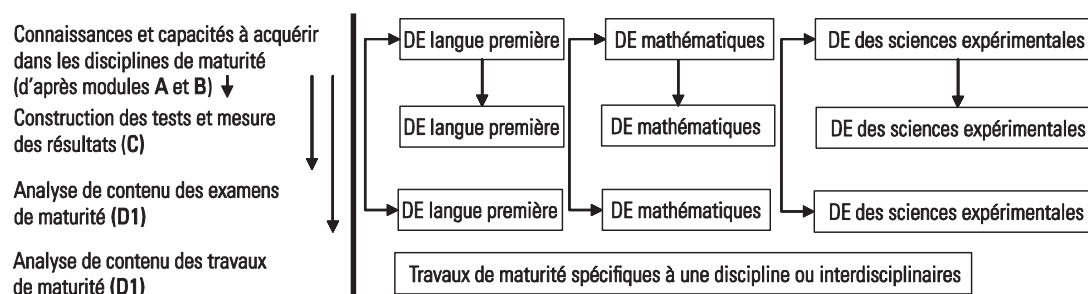
Pour plusieurs raisons (pratiques, non pas fondamentales), il n'était pas possible de réaliser toutes les enquêtes esquissées sous (2) dans le temps à disposition. C'est pourquoi il a fallu effectuer une sélection pour EVAMAR II. En regard du modèle global, il s'agit d'une variante minimale, mais qui du moins était encore finançable. La figure I.2 présente le résultat de cette sélection.

Figure I.2: Enquêtes choisies pour EVAMAR II

EVAMAR II, 1^{re} partie: Définition des critères d'évaluation (modules A et B)



EVAMAR II, 2^e partie: Procédure d'évaluation (modules C et D)



Légende: - Les lettres A à D désignent les différents modules du projet.
- DE: abréviation de «discipline enseignée»

Dans le cadre d'EVAMAR II, il était prévu de définir concrètement les objectifs de la formation de maturité pour des éléments choisis du premier objectif principal, l'aptitude aux études supérieures. Le deuxième objectif principal (celui d'une profonde maturité sociale grâce à une vaste culture générale) n'a donc été que partiellement intégré dans l'opérationnalisation ciblée choisie par le mandant (dans le domaine de recoupement) et devrait être traité dans d'autres projets (en dehors d'EVAMAR II).

Pour modéliser de façon opérationnelle la notion abstraite d'aptitude aux études supérieures, il était primordial d'identifier aussi bien les connaissances, aptitudes, savoir-faire et dispositions interdisciplinaires utiles à des études réussies que les compétences nécessaires dans chaque discipline. La question centrale était de connaître les exigences posées aux étudiants, dans leurs études et leur travail au niveau des hautes écoles, en termes de processus effectifs d'apprentissage et de pensée, et de déterminer quelles compétences spécifiques et

interdisciplinaires sont requises dans ce cadre. A cet effet, on a retenu pour EVAMAR II les enquêtes suivantes, issues du concept global décrit à la section 2.

- A. Enquêtes visant à clarifier, du point de vue des universités, le concept de «maturité requise pour entreprendre des études supérieures» ou «aptitude aux études supérieures».
 - A1. Analyse des contenus des cursus, des notes de cours, des manuels principaux, etc. dans les universités: détermination des prérequis spécifiques aux différentes disciplines d'études.
 - A2. Analyse des examens d'admission pertinents: détermination des exigences cognitives.
 - A3. Analyse des contenus des premiers examens dans les universités (examens intermédiaires; examens des niveaux «assessment»): détermination des exigences cognitives.
 - A4. Enquête complémentaire auprès des enseignants chargés des cours dans le cadre des semestres de base quant aux exigences et aux prérequis en matière d'études.
- B. Objectif opérationnalisé de la formation de maturité quant à l'«aptitude aux études supérieures»: développement de grilles de compétences visant à systématiser les résultats des enquêtes précédentes.

Les enquêtes A et B devaient ainsi déboucher sur les composantes essentielles des exigences effectivement posées aux étudiants au début de leurs études. Il faut noter, s'agissant de ces enquêtes visant à définir les critères d'évaluation de l'«aptitude aux études supérieures», que les résultats ne correspondent pas nécessairement à un optimum déterminé, même seulement recherché, car les exigences déterminées sont susceptibles de comporter déjà l'adaptation des étudiants à des conditions non optimales. En outre, les exigences des hautes écoles se trouvent actuellement elles-mêmes dans un processus dynamique (réforme de Bologne). Les exigences effectivement posées aux étudiants par les hautes écoles universitaires devraient donc être également périodiquement redéfinies (du haut vers le bas), dans le cadre d'un suivi sur le long terme après EVAMAR II, et soumises à un discours critique (du bas vers le haut). Dans ce contexte, il serait alors éventuellement possible de conduire également une discussion sur les standards de formation au gymnase.

S'agissant de l'enquête concernant l'objectif de l'«aptitude aux études supérieures», les modules suivants ont été choisis pour EVAMAR II à partir du concept global (cf. section 5.1).

- C. Résultats de la formation de maturité: le module C, qui prend appui sur les résultats des modules A et B, se concentre sur le développement et l'exécution de tests visant à mesurer les aspects des résultats de la formation de maturité pertinents pour les hautes écoles dans les trois disciplines de la langue première, des mathématiques et de la biologie. Ces résultats ont été complétés, pour les raisons exposées ci-après (cf. chapitre IV), par une enquête supplémentaire et (dans un sous-groupe de gymnases) par une partie de test permettant de saisir les aptitudes cognitives générales pertinentes pour les hautes écoles. En outre, les évaluations ont tenu compte des notes de maturité effectivement réalisées.

De ce fait, par rapport au concept global, on a réduit le nombre de disciplines pour lesquelles devaient être élaborés et conduits des tests visant à mesurer la performance. On a retenu les mathématiques, la langue première et une branche scientifique expérimentale à titre de variante minimale. La langue première et les mathématiques sont désignées avec l'anglais, dans la littérature correspondante, comme les disciplines dans lesquelles sont acquises entre autres des aptitudes de base importantes pour toutes les disciplines d'études (par ex. Huber, 1997 et 1998). Le fait que l'on ait retenu une branche scientifique expérimentale plutôt que l'anglais, qui était proposé par le mandataire, relève de la décision du mandant. Les sciences expérimentales se situent de fait directement au centre de la discussion de politique de la formation concernant le RRM de 1995. Le choix des branches retenues correspond donc aussi à un consensus politique des mandants. La biologie a été choisie ultérieurement comme discipline scientifique expérimentale (on trouvera une justification de ce choix au chapitre IV).

- D. Qualité des examens de maturité et du travail de maturité: le module D doit permettre d'étudier les aspects des mesures de la performance réalisée en interne dans les gymnases au moyen des examens de maturité. A cet égard, il faut noter que les notes de maturité ne sont que les indicateurs du dernier maillon «examens de

maturité» dans la chaîne des qualifications acquises au gymnase et qu'elles ne revêtent à ce titre qu'une signification limitée. Elles ne restituent pas nécessairement les compétences effectives des bacheliers et des bachelières. Il est connu que de nombreux facteurs exercent une influence entre les objectifs fixés par l'art. 5 RRM et le PEC, d'une part, les qualités interprétables à partir des notes obtenues lors des examens de maturité, d'autre part. Des enquêtes correspondantes permettraient de relever ces facteurs et de les mettre en relation l'un avec l'autre. Il s'agit par exemple des facteurs suivants: (1) la mise en œuvre des directives du RRM de 1995 dans les grilles et plans d'études locaux; (2) la qualité de l'enseignement et (3) la qualité des examens de maturité et des travaux de maturité. Pour des raisons liées aux ressources, les enquêtes d'EVAMAR II se sont limitées au troisième aspect. Les sous-projets suivants ont ainsi été réalisés.

- D1. Analyse des contenus d'examens de maturité (matière spécifique par disciplines et niveau des exigences cognitives) et comparaison des résultats avec les compétences cibles déterminées dans le module B, compte tenu des résultats de la mesure centrale des performances dans le module C.
- D2. Analyse de travaux de maturité: examen visant à déterminer dans quelle mesure les élèves de maturité ont appris à travailler scientifiquement grâce au travail de maturité; on a procédé à cet examen en évaluant les travaux a posteriori (évaluation d'experts), en interrogeant les auteurs des travaux de maturité visés et en reconstruisant les conditions institutionnelles cadre.

5.3 Limites de l'étude EVAMAR II

En comparant le concept global (section 5.1) avec les parties choisies qui en sont issues (section 5.2), il apparaît qu'EVAMAR II ne constitue pas une évaluation complète de la formation gymnasiale: fondamentalement, le projet ne permet de s'exprimer que sur quelques aspects choisis des objectifs et de la réalisation des objectifs de la formation gymnasiale et il ne livre que des observations de points de vue sélectionnés de l'interface entre le gymnase et l'université.

D'abord, il ne s'agit que de connaissances et de capacités qu'il est possible de saisir par des tests de produits d'apprentissage et des tests d'aptitude, ainsi que par auto-évaluation. Les connaissances et les capacités correspondant à des compétences globales au sens de Weinert (2001) ne sauraient être mesurées dans le cadre du projet pour des raisons méthodologiques et de ressources. Nous songeons en l'occurrence, par exemple, aux aptitudes dans le domaine de l'acquisition des connaissances (conception d'hypothèses, techniques de travail pratiques comme les techniques d'expérimentation, interprétation subséquente des données, etc.), aux capacités de communication écrite et orale, à la capacité d'évaluer des faits complexes sur la base de critères et aux aptitudes créatrices en général. Dans ce contexte, relevons toutefois que, sous l'angle de la psychologie de la cognition, les connaissances spécifiques aux disciplines qui fondent la capacité à comprendre, à poursuivre l'apprentissage et à se remémorer rapidement revêtent une importance beaucoup plus grande que ne le suggère parfois la vaste discussion en matière de compétences. Aujourd'hui, les avantages donnés par de solides connaissances spécifiques aux différentes disciplines pour comprendre, apprendre et se souvenir sont expérimentalement bien établis, justement en ce qui concerne les textes spécialisés (par ex. Weinert, 2001; cf. présentation détaillée au chapitre IV, section 1.2.1). La qualité des connaissances spécifiques atteinte dans les différentes disciplines enseignées au gymnase est particulièrement pertinente pour EVAMAR II dans la perspective pragmatique – mais en aucune manière négligeable – de l'aptitude générale aux études; ce d'autant que le projet se concentre sur l'interface gymnase-université, c'est-à-dire en particulier sur l'aptitude à étudier tout au début des études. S'agissant de textes spécialisés, les aptitudes de prime abord «banales» consistant à comprendre, à apprendre et à se souvenir semblent d'une importance plus déterminante pour franchir au moins les tout premiers obstacles d'examens universitaires que la qualité d'aptitudes plus axées sur l'action (et la profession), comme l'acquisition des connaissances, l'aptitude à communiquer et la capacité de jugement.

Compte tenu des limites imposées au projet en termes de méthodes et de ressources, on ne peut par ailleurs prétendre couvrir l'ensemble des connaissances et capacités spécifiques exigées dans les universités, mais

seulement celles que l'on aura pu identifier à partir des supports d'enseignement et d'examen disponibles en se fondant sur l'appréciation des experts engagés pour l'analyse de ces derniers. Dans la perspective de l'utilisation dans le module C (développement et réalisation des tests) des résultats obtenus dans le module A (analyse des supports), la méthodologie imposait de procéder à une sélection de ces volumes, même limités, de données partielles, sans quoi les tests auraient été beaucoup trop vastes. Il s'agit cependant d'une sélection considérable des connaissances et des capacités initiales effectivement nécessaires pour entreprendre des études de son choix.

Le projet EVAMAR II n'examine donc pas les aspects, énumérés ci-après, que comporterait un tableau complet de l'aptitude globale aux études supérieures et d'une profonde maturité sociale; le projet n'analyse donc pas si les bacheliers et les bachelières en disposent.

- Les connaissances et les capacités, qui font certes partie d'une vaste culture générale (maturité sociale constituant la base voulue pour assumer des responsabilités dans la société) et qui doivent entrer à ce titre dans la formation gymnasiale, mais qui ne sont pas directement requises à l'université.
- Les connaissances initiales nécessaires dans les branches d'études qui n'ont pas été sélectionnées (cf. chapitre II, section 1).
- Les connaissances et les capacités nécessaires à l'aptitude aux études supérieures qui sont inhérentes à d'autres disciplines d'enseignement que la langue première, les mathématiques et la biologie.
- Les éléments de compétences interdisciplinaires que l'on n'aurait pu saisir que difficilement dans le cadre donné du projet, mais qui sont toutefois importants pour gérer fructueusement des études (par ex. la capacité à organiser soi-même son apprentissage, à s'organiser soi-même en général, à effectuer des recherches, à traiter rapidement une vaste littérature spécialisée, etc.). Seules des enquêtes supplémentaires, proposées dans le concept global, permettraient de déterminer l'importance de tels facteurs (enquêtes longitudinales parmi les étudiants) et seules de lourdes méthodes d'évaluation permettraient de les mesurer. Toujours est-il que le projet EVAMAR I (cf. section 2) et d'autres études d'évaluation (cf. section 3) ont déjà permis de valider approximativement ces facteurs par des enquêtes, autrement dit par le biais d'auto-évaluations. En outre, on a inséré des questions complémentaires sur EVAMAR I dans le questionnaire supplémentaire d'EVAMAR II.

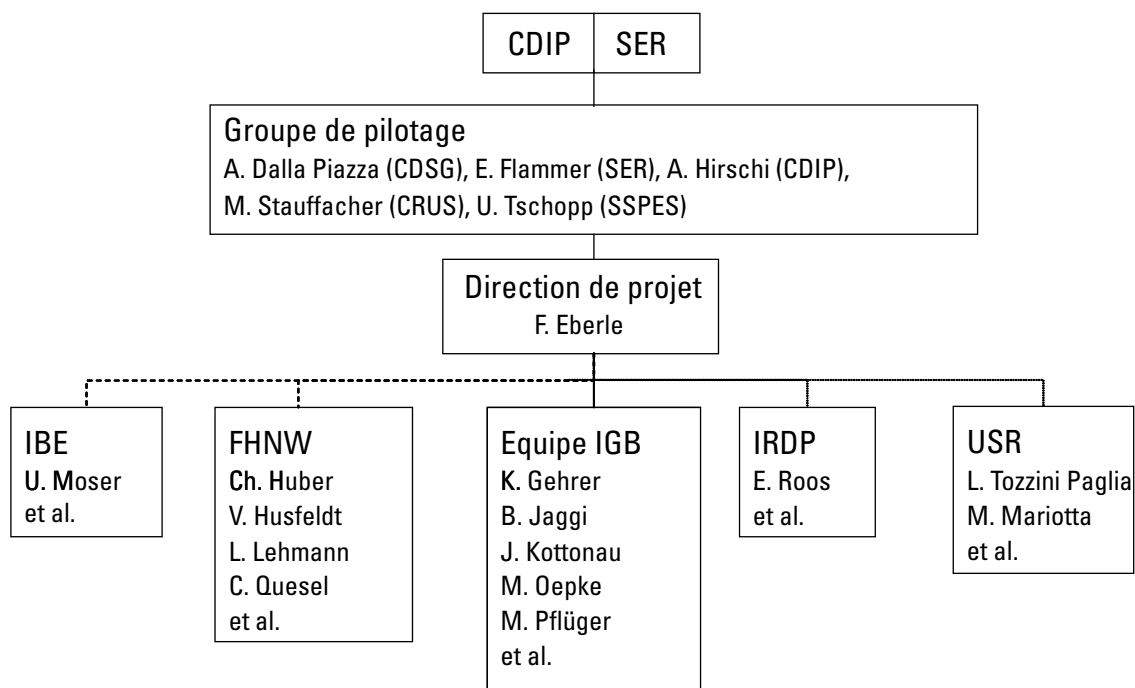
Ainsi, en se fondant sur les enquêtes choisies pour EVAMAR II, on ne peut s'exprimer que sur des aspects *partiels* de l'aptitude aux études supérieures. De plus, l'objectif de préparer à «assumer des responsabilités au sein de la société» n'a pas été directement étudié. Il faudra tenir compte de ces restrictions au moment d'interpréter les résultats et d'en déduire les mesures visant à façonner le gymnase de l'avenir.

Dans la discussion sur le projet EVAMAR II, on a souvent entendu dire que les standards de formation futurs seraient fixés pour le gymnase dans les modules A et B. Il semble qu'il s'agisse d'un malentendu. Les modules A et B sont voués à une analyse de la situation actuelle d'une partie de l'interface gymnase-université. Leurs résultats ne sauraient revêtir le statut d'objectifs. Pour fixer les exigences à atteindre, il faudrait, entre les institutions participant à l'interface, des processus de coordination supplémentaires, étayés aussi largement que possible, qui ne pourraient déboucher qu'ensuite sur des standards de formation assumés conjointement par les universités et les gymnases. Alors seulement, il s'agirait de futures exigences de performance officielles que les organismes compétents auraient reformulées et décidées de manière contraignante. En déterminant de telles exigences, on devrait par ailleurs discuter dans toute leur étendue les objectifs de la formation gymnasiale prévus à l'art. 5 du RRM de 1995. Pour ces raisons, les résultats d'EVAMAR II ne peuvent constituer que partiellement la base d'éventuels travaux visant le développement des standards de formation.

5.4 Organisation de projet et remerciements

La structure organisationnelle du projet apparaît dans l'organigramme de la figure I.3.

Figure I.3: Organigramme du projet EVAMAR II



Légende: - CDIP: Conférence suisse des directeurs cantonaux de l'instruction publique
- SER: Secrétariat d'Etat à l'éducation et à la recherche
- CDSG: Conférence des directeurs des gymnases suisses
- CRUS: Conférence des recteurs des universités suisses
- SSPES: Société suisse des professeurs de l'enseignement secondaire
- IBE: Institut für Bildungsevaluation, Université de Zurich
- IGB: Institut für Gymnasial- und Berufspädagogik, Université de Zurich
- FHNW: Haute école spécialisée du Nord-Ouest de la Suisse, Haute école pédagogique, Institut Forschung und Entwicklung
- IRDP: Institut de recherche et de documentation pédagogique, Neuchâtel
- USR: Ufficio studi e ricerca, Bellinzona

Le mandat de direction du projet EVAMAR II a été confié au professeur Franz Eberle de l'Institut de pédagogie gymnasiale et professionnelle IGB de l'Université de Zurich (Institut für Gymnasial- und Berufspädagogik, anciennement Höheres Lehramt Mittelschulen).

L'équipe centrale de l'IGB comprenait les collaboratrices et collaborateurs scientifiques suivants: Nina Barske (jusqu'à novembre 2007), Karin Gehrler (langue première), Daniel Halter (jusqu'à juin 2006), Beat Jaggi (d'août 2006 à juillet 2008, mathématiques), Johannes Kottonau (dès juillet 2006, biologie), Maren Oepke (direction opérationnelle, langue première) et Michael Pflüger. Les indications entre parenthèses renvoient à certaines fonctions ou compétences spécifiques assumées dans le cadre du projet. Nombre d'autres collaborateurs scientifiques ont été engagés temporairement dans le cadre du projet et affectés à d'autres tâches spécifiques. Citons notamment Dirk Bauer, Christine van Bedaf et Claudia Illés. Ruth Berri, du secrétariat de la Chaire, a accompli de larges parts des travaux administratifs.

En ce qui concerne la Suisse romande, nos partenaires de coopération de l'Institut de recherche et de documentation pédagogique (IRDP) de Neuchâtel, soit Matthis Behrens, directeur, et Eva Roos, collaboratrice scientifique, ont assumé plusieurs tâches, notamment en contribuant au développement et à l'adaptation des instruments, en assurant le relevé des données en Suisse occidentale et en participant à l'évaluation des données.

Au Tessin, nous avons reçu l'appui des partenaires de l'Ufficio studi e ricerca (USR) de Bellinzone, à savoir Emanuele Berger, directeur, Luana Tozzini Paglia et Myrta Mariotta.

Urs Moser et des collaborateurs de l'Institut pour l'évaluation de la formation IBE (Institut für Bildungsevaluation), associé à l'Université de Zurich, ont contribué de manière déterminante aux tests du module C, notamment à l'échantillonnage, à la préparation des pré-tests et à la modélisation des données selon le modèle Rasch.

Pour le test d'aptitudes interdisciplinaires, nous avons coopéré avec le professeur Klaus-Dieter Hänsgen du Centre pour le développement de tests et le diagnostic (CTD) de l'Université de Fribourg.

Le sous-projet «Analyse des travaux de maturité» (D2) a été intégralement conçu et réalisé par la Haute école pédagogique de la HES de la Suisse du Nord-Ouest (Christina Huber, professeur Vera Husfeldt, Lukas Lehmann et professeur Carsten Quesel [direction]). Cette équipe a été renforcée par Felix Boller (HEP/HES du Nord-Ouest de la Suisse) et par Eva Roos (IRDP).

A toutes les personnes citées, nous voulons adresser nos vifs remerciements. Nous remercions en particulier la CDIP et le SER, nos mandants, de la confiance témoignée et tous les membres du groupe de pilotage pour leur encadrement extrêmement constructif du projet. Notre gratitude va également à tous les bacheliers et bachelières qui se sont montrés disposés à passer des tests et à remplir les questionnaires ou encore à nous remettre leurs travaux de maturité en vue de l'évaluation.

Nous sommes en outre reconnaissants aux nombreux membres du corps enseignant, aux membres des directions d'école et au personnel administratif, qui se sont employés à nous aider lors du relevé des données. Enfin, nous souhaitons remercier la Conférence des Directrices et directeurs de gymnases suisses (CDGS) et la Société suisse des professeurs de l'enseignement secondaire (SSPES) de nous avoir soutenus moralement.

II Module A: éléments de l'aptitude aux études selon les hautes écoles

Les sous-projets A1 à A4 devaient permettre, dans une première étape principale de l'évaluation (définition des critères d'évaluation, c'est-à-dire des objectifs de la formation de maturité), de déterminer les connaissances et les capacités requises pour entrer dans une université suisse. Après seulement, dans une deuxième étape principale de la procédure d'évaluation, il devait être possible d'analyser la réalisation de ces objectifs en définissant toute une variété d'indicateurs. On a pu déduire les objectifs de ceux que la loi fixe – sous une forme générale seulement – aux écoles de maturité. Les principaux objectifs de la formation gymnasiale sont, selon l'art. 5 du RRM de 1995, de conférer d'une part «la maturité requise pour entreprendre des études supérieures» (maturité nécessaire ou aptitude aux études supérieures) et de préparer d'autre part «à assumer des responsabilités au sein de la société actuelle» (profonde maturité sociale grâce à une vaste culture générale). Ces objectifs ne se recouvrent que partiellement. Dans le cadre du projet EVAMAR II, le mandat ne visait la concrétisation des objectifs de la formation de maturité que pour des aspects choisis de l'objectif principal d'aptitude aux études. Le deuxième objectif principal (profonde maturité sociale grâce à une vaste culture générale) n'est donc intégré dans la définition opérationnelle du but visé que pour le domaine où les deux objectifs cités se recoupent (cf. détails au chapitre I, section 5). A cet effet, diverses méthodes sont utilisées dans les sous-projets. Nous présentons ci-après en détail la base des données et les différents sous-projets.

1 Base de données des sous-projets A1, A3 et A4

Pour des raisons de ressources, il n'a pas été possible de procéder à un relevé complet. L'enquête a couvert les 16 disciplines les plus fréquemment choisies (selon le nombre d'étudiants), correspondant à environ 70% de tous les étudiants immatriculés en Suisse pendant l'année académique 2004/2005. L'enquête de l'Office fédéral de la statistique concernant les étudiants de l'année 2004/2005 (OFS, 2008a) a fourni la base de données à cette sélection. Par ailleurs, la question s'est posée de savoir s'il serait vraiment judicieux (également compte tenu du coût) d'analyser plusieurs fois des filières d'études comparables, alors que des résultats semblables étaient probables. C'est pourquoi le choix des disciplines à examiner s'est concentré sur les deux ou trois universités où étaient enregistrés cumulativement plus de 50% de tous les étudiants de la branche visée. En raison de cette sélection à deux niveaux, environ 37% de tous les étudiants immatriculés en Suisse ont été couverts par cette sélection des filières d'études. Le tableau II.1 présente les disciplines d'études choisies, par ordre de taille des populations d'étudiants correspondantes, ainsi que les universités choisies et prises en compte selon la méthode citée. Cette sélection a constitué la base de données pour les enquêtes des sous-projets A1, A3 et A4.

Tableau II.1: Disciplines d'études retenues dans la base de données du module A

Orientation des études	Universités
Economie	Université de Zurich, Université de Saint-Gall
Droit	Université de Zurich, Université de Fribourg
Psychologie	Université de Zurich, Université de Berne
Médecine humaine	Université de Zurich, Université de Lausanne
Histoire	Université de Zurich, Université de Berne
Biologie	Université de Zurich, EPFZ, Université de Bâle
Sciences politiques	Université de Genève, Université de Zurich
Architecture & planification	EPFZ, Université de Lugano
Informatique	EPFZ, EPFL
Sociologie	Université de Fribourg, Université de Zurich
Germanistique	Université de Zurich, Université de Bâle
Communication & médias	Université de Zurich, Université de Lugano
Français	Université de Lausanne, Université de Genève
Sciences de l'éducation	Université de Genève, Université de Zurich
Anglais	Université de Zurich, Université de Berne
Physique	EPFZ, EPFL

2 Sous-projets A1 et A3: analyse de contenu des supports d'études et des examens

Dans le cadre du SP A1, on a soumis à une analyse de contenu les cursus, les notes de cours et les principaux manuels des différentes disciplines au cours des deux premiers semestres d'études. Il s'agissait de déterminer les connaissances et les capacités requises pour comprendre et traiter les supports d'enseignement. A cet effet, on a développé une grille d'analyse et un manuel d'encodage correspondant, qui ont servi aux experts engagés dans le projet de guide au codage des supports. S'agissant des critères d'analyse, il fallait chercher ceux qui semblaient également applicables à l'analyse des supports d'études très différents de toutes les disciplines. Le concept de connaissance de Bloom et al. (1956 et 1974), non spécifique aux disciplines, dans la version soigneusement remaniée et élargie par Anderson et al. (2001), au sens d'une adaptation aux éléments nouveaux de la psychologie de la cognition, est apparu fournir une base appropriée à la constitution des catégories recherchées. Il permet des commentaires comparatifs sur toutes les disciplines et il s'agit de l'un des modèles pédagogiques-didactiques les mieux éprouvés et les plus répandus de ces dernières décennies (cf. les explications à ce sujet au chapitre V, section 1.2).

Comme les supports d'enseignement comprennent souvent des exercices et des tâches dont le traitement requiert des étudiants des aptitudes cognitives qui vont au-delà de la pure réception de connaissances, on a développé pour le codage des tâches un système de catégories, adapté à partir du codage des textes didactiques, qui intègre ces aptitudes cognitives. Il est finalement apparu que le codage des tâches pouvait être repris pour l'analyse des premiers examens du cursus d'études à réaliser dans le sous-projet A3. La description qui suit de la méthode se rapporte donc aux deux sous-projets A1 et A3, puisqu'ils reposent sur la même base de données.

2.1 Méthode et instruments du SP A1 et du SP A3

La méthodologie appliquée dans EVAMAR II, dans les projets A (et B) est inédite, si bien qu'il n'était pas possible de recourir à des instruments déjà existants. Pour analyser les supports, on a développé deux procédures de codage: le «codage fin» et le «codage large».

2.1.1 Méthode du codage fin

Avec ce type de codage, on distingue entre le codage du contenu des supports d'enseignement, d'une part, et le codage des tâches ou des exercices qui s'y rapportent et (comme déjà décrit) des examens (SP A3), d'autre part.

S'agissant du contenu des supports d'enseignement, on a déterminé des unités de sens que les étudiants doivent déjà connaître pour comprendre le matériel en question. Les unités de sens peuvent être formées de mots déterminés particuliers (noms, adjectifs, adverbes, etc.), parties de phrase, de phrases entières et/ou de plusieurs phrases du texte original et apparaître elles-mêmes sous ces formes. Ces unités de sens supposées acquises dans le texte didactique ont été classées en diverses catégories de connaissances, axées pour l'essentiel dans la première dimension sur les catégories de Bloom (1956 et 1974) ou d'Anderson et al. (2001) et, dans la deuxième dimension, sur les disciplines de maturité. Une même unité de sens peut donc être attribuée à plusieurs disciplines scolaires.

Le tableau II.2 représente sous une forme simplifiée la grille prescrite, dans laquelle les unités de sens sont venues s'insérer. Les lignes hébergent les catégories de connaissance (AA à DB, cf. ci-après pour des explications), tandis que les colonnes contiennent les abréviations pour les disciplines scolaires du gymnase (toutes ces disciplines n'étant pas enseignées dans tous les gymnases). Cette figure montre que, outre les domaines de connaissance usuels – allemand (All), français (F), italien (I), anglais (Ang), latin (L), grec (Gr), mathématiques (M), biologie (B), chimie (C), physique (P), histoire (H), géographie (G), économie et droit (ED), arts visuels (Av), musique (Mu) et philosophie, pédagogie, psychologie, éthique et enseignement religieux (PPP), on a saisi des catégories interdisciplinaires telles que les travaux scientifiques (Ts, par ex. types de travaux

écrits, planification du temps, théorie de la connaissance) et l'informatique (I). La dernière colonne (Div) donnait la possibilité de saisir des unités de sens qui n'entraient pas dans les autres catégories scolaires typiques (par ex. les symboles).

Tableau II.2: Grille prescrite pour saisir la dimension cognitive et les disciplines scolaires

	All	F	I	Ang	L	Gr	M	B	C	P	H	G	ED	Av	Mu	PPP	Ts	Inf	Div
AA																			
AB																			
BA	US1																		
BB	US2																		
BC																			
CA																			
CB	US3																		
CC																			
DA																			
DB																			

Légende:

- US = unité de sens
- lettres en abscisse: abréviations des disciplines d'enseignement gymnasiales (Ts = travaux scientifiques, Inf = informatique, Div = divers)
- lettres en ordonnée: abréviations pour les catégories de connaissance

Outre leur attribution aux disciplines d'enseignement gymnasiales mentionnées en colonnes, les unités de sens ont été évaluées quant à leur attribution possible, en fonction de leur emplacement dans le texte, aux catégories de connaissance présentées au tableau II.3 (en lignes): connaissances factuelles (AA et AB), connaissances conceptuelles (BA à BC), connaissances procédurales (CA à CC) ou connaissances métacognitives (DA et DB). On trouvera une description détaillée et une classification plus fine de ces catégories de connaissance appliquées au codage dans le manuel de codage, qui comporte également des exemples (cf. rapport technique).

Tableau II.3: Catégories de connaissance des unités de sens (catégories principales)

-
- A. Connaissances factuelles
 - AA Connaissance des terminologies
 - AB. Connaissance des détails et éléments généraux
 - B. Connaissances conceptuelles
 - BA. Connaissance des classifications et des catégories
 - BB. Connaissance des principes et des généralisations
 - BC. Connaissance des théories, des modèles et des structures
 - C. Connaissances procédurales
 - CA. Connaissance des qualifications spécifiques (à la discipline)
 - CB. Connaissance des techniques et de méthodes spécifiques (à la discipline)
 - CC. Connaissance des critères de détermination d'une procédure appropriée
 - D. Connaissances métacognitives
 - DA. Connaissance de la stratégie
 - DB Connaissance des tâches cognitives, y compris la connaissance du contexte et des conditions

Ces connaissances ont également été saisies lors du codage des tâches, des exercices et des examens (SP A3), si elles constituaient des prérequis à la solution et des connaissances d'entrée dans le cursus d'études, aucune explication ne se trouvant dans les documents d'enseignement. La classification s'est alors limitée aux catégories de connaissance principale A à D. En revanche, si une tâche ou un examen faisait appel uniquement à des connaissances préalablement transmises dans le matériel d'enseignement, on ne procédait à aucune saisie sur le plan des connaissances.

Outre la saisie au cas par cas de la catégorie de connaissance, on a noté pour chaque tâche, chaque exercice et chaque examen les aptitudes cognitives nécessaires à résoudre la tâche en question. La catégorisation de ces aptitudes cognitives est également axée sur la taxonomie de Bloom (1956 et 1974, développée par Anderson et al., 2001). Les encodeurs ont saisi les processus de pensée à suivre pour résoudre la tâche, comme lors de l'analyse des contenus d'enseignement: à l'aide d'une grille correspondante. La catégorisation et l'attribution des aptitudes cognitives sont représentées au tableau II.4. On trouvera également dans le manuel de codage une description détaillée et des exemples (cf. rapport technique).

Tableau II.4: Catégories d'aptitudes cognitives pour les tâches et les examens

- A. Se remémorer.
- B. Comprendre. Les exigences suivantes sont comprises sous ce point: interpréter (clarifier, paraphraser, traduire), expliquer par l'exemple (illustrer), classer (catégoriser, subsumer), résumer (abstraire, généraliser), déduire (conclure, prédire, extrapoler), comparer (contraster), expliquer (modéliser).
- C. Appliquer. Comprenons les actions suivantes: traiter (exécuter), mettre en œuvre (appliquer), réfléchir simplement à l'application propre.
- D. Analyser. Cette catégorie comprend: différencier (distinguer, cibler, sélectionner), organiser (identifier, montrer les corrélations/la cohérence, intégrer, esquisser, structurer, déterminer la syntaxe/grammaire), attribuer (déterminer les composantes, établir les liens de causalité, classer), transposer à une situation réelle (par ex. traduire la situation dans la langue des mathématiques, de la physique, etc., lui trouver une solution dans cet univers, interpréter le résultat pour la situation réelle).
- E. Evaluer. On entend les activités suivantes: vérifier (calculer, coordonner, contrôler, surveiller, tester), soumettre à l'appréciation critique (juger).
- F. Créer. Comprenons par ce terme: générer (établir des hypothèses, abstraire, combiner les éléments de réflexion pour préparer une conclusion), planifier (projeter), produire (construire), composer (au sens de découvrir ou inventer des relations/structures).

Des critères déterminés ont permis de décrire et d'évaluer les figures et les graphiques, tout en offrant la possibilité de coder les catégories ouvertes, si bien qu'on a pu procéder à une saisie, lorsque d'autres compétences, comme les compétences rhétoriques ou motrices, étaient nécessaires aux étudiants pour comprendre les textes ou pour résoudre les tâches (cf. à cet égard les détails du manuel dans le rapport technique).

Le «codage fin» représente donc globalement une méthode de codage très précis et détaillé.

Le système de codage a été développé avec le concours des encodeurs et il a fait l'objet de plusieurs discussions. Tous les encodeurs, qui sont également des experts dans la discipline d'études qu'ils doivent analyser, ont suivi une formation adéquate. Aux fins de contrôle de fiabilité, trois mesures différentes ont été calculées pour établir la fidélité inter-juges: la concordance (par paires) dans le choix des unités de sens, dans l'attribution aux catégories de connaissance et dans l'attribution aux disciplines. Nous n'avons obtenu que des valeurs moyennes ($r = .5$ pour le choix des unités de sens). Elles ont donné lieu à des discussions et à des compléments de formation en vue d'améliorer l'homogénéité des codages. On trouvera ci-après une appréciation de ces valeurs en rapport à leur contexte (cf. section 2.1.3).

2.1.2 Méthode du codage large

La méthode du codage fin, qui prend beaucoup de temps, a été appliquée pour la saisie intensive et précise des détails des chapitres introductifs de manuels ou de notes de cours. Pour analyser des modèles/théories entiers et, généralement, des unités de sens dans un contexte plus étendu, on a développé une méthode de codage large pour les chapitres suivants de ces mêmes manuels, car elle est moins gourmande en temps. Tandis que le codage des premiers chapitres des supports d'enseignement introductifs et des exercices qui leur sont associés était réalisé à l'aide de la méthode fine décrite ci-dessus, les supports didactiques suivants ont été soumis à un processus d'analyse plus rapidement exécutable: la méthode du «codage large». Il s'agissait toujours de chercher des unités de sens que les étudiants devaient déjà connaître pour pouvoir comprendre le matériel d'enseignement (connaissances initiales ou de départ). Mais le codage large ne porte plus seulement sur certains mots, parties de phrases et/ou phrases formulant ces connaissances de départ, il peut également concerner des

thèmes entiers issus des disciplines scolaires sous-jacentes (par ex. «mathématiques: dériver des fonctions dans le cadre du calcul différentiel»). Dans la méthode dite «large», contrairement au codage «fin», on se concentre sur l'analyse des unités de sens supposées connues pour que le message central de l'enseignement, les éléments centraux des chapitres, soient compréhensibles. A cet égard, on a pris en compte que la connaissance de l'ensemble du texte ou de chacune des notions n'est pas toujours nécessaire pour comprendre un message textuel. Ainsi, des passages entiers du texte peuvent, dans certains cas, rester incompris sans que la compréhension du message central de l'enseignement n'en pâtisse nécessairement. Les connaissances à catégoriser selon cette méthode de détermination ont également été attribuées aux disciplines scolaires et saisies dans des grilles correspondantes.

2.1.3 Qualité des techniques de mesure

Comme la présente étude repose sur une méthode qualitative, il était en outre nécessaire de conduire des réflexions spécifiques quant à la qualité des techniques de mesure. Commençons par les questions de fiabilité et de validité.

- En ce qui concerne la fiabilité, on aurait éventuellement pu accroître les valeurs de la fiabilité, moyennes seulement, en effectuant un deuxième, voire un troisième codage a posteriori. Cette optimisation n'était pas possible dans ce projet en raison des ressources disponibles. Si la concordance ne reçoit que des valeurs moyennes, c'est qu'elles résultent du caractère subjectif de l'estimation des encodeurs qui déterminent les connaissances et capacités supposées acquises dans les textes étudiés. Il est en particulier difficile de distinguer clairement les unités de sens selon le critère de savoir s'il s'agit de culture générale allant de soi ou de connaissances (gymnasiales) exigeantes. Contrairement aux connaissances exigeantes, la culture générale ne devait pas être saisie, car même si elle est également indispensable, elle ne fait pas partie des connaissances sélectives dans la perspective de la réussite des études et, généralement, parce qu'elle est acquise en dehors du gymnase. Si l'on se réfère au modèle d'une pyramide des connaissances, la culture générale qui va de soi se situe dans les parties inférieure et moyenne, tandis que les connaissances exigeantes se trouvent dans la partie supérieure. Dès lors se pose la question des effets sur les objectifs du projet de cette fidélité inter-juges dont les valeurs ne sont que moyennes. Même si les encodeurs n'ont pas attribué le même nombre d'unités de sens dans le domaine moins exigeant, on observe une forte concordance en ce qui concerne les unités de sens présentant une certaine difficulté, qui sont plutôt pertinentes pour l'application ultérieure. Il fallait de toute façon procéder à une sélection à partir de l'univers d'unités de sens obtenu pour élaborer les tests (module C). Que l'on ait axé cette opération sur les unités de sens exigeantes se justifie par le fait que celles-ci représentent le véritable obstacle à surmonter pour conduire fructueusement des études. La sélection des connaissances et des capacités, dans les projets B et C, reposent donc sur les unités de sens exigeantes présentant une fidélité inter-juges supérieure lors de leur identification.
- Quant à la validité des tests à développer, elle ne souffre également pas de la méthodologie de codage adoptée, conformément aux explications précédentes. Dans ce contexte, notons que les unités de sens à choisir pour les tests ne constituent qu'une partie des connaissances requises pour mener des études universitaires. Mais cette sélection ne contient pas non plus d'élément dont la nécessité pour des études ne serait pas prouvée ou qui n'y participerait pas à tout le moins de manière avérée. Toutefois, le résultat du codage fin consiste en diverses unités de sens relativement dissociées l'une de l'autre, qui ne suffisent pas à fournir un tableau cohérent des connaissances et des capacités nécessaires à des études déterminées. A cet effet, il fallait encore préparer les données dans leur contexte d'application (module C). Le problème posé par la présence d'unités de sens microscopiques, pour certaines dissociées de leur contexte sémantique général, est atténué par la méthode de codage large, qui permet plutôt de saisir des thèmes globaux. Cependant, le résultat des analyses n'est évidemment pas encore assimilable à «l'aptitude aux études».

Assurément, la compréhension de certains concept suppose souvent une reconstruction du sens pendant le processus même de lecture, qui suppose l'aptitude à comprendre des notions inconnues ou incomprises jusque

là à partir du contexte sémantique. Or, dans la logique de notre évaluation, ces notions constituent clairement des connaissances de départ. A cette objection, on peut répondre que les concepts en question ont déjà dû être appris une fois en une occasion déterminée. Si toute trace de souvenir fait défaut, la compréhension n'est pas correcte. Evidemment, l'étudiant peut combler de telles lacunes de connaissance de cas en cas par une recherche. Mais s'il doit excessivement procéder à de telles investigations, il éprouve des difficultés considérables dans ses études. L'enquête repose donc sur le type de l'étudiant «idéal», qui possède dans son bagage toutes les connaissances préalables requises dès le début de ses études. Cette argumentation est considérablement renforcée par le fait, déjà relevé précédemment, que les connaissances propres à une discipline dont l'étudiant dispose constituent une base importante de son apprentissage à venir (par ex. Weinert, 2001; les éléments théoriques détaillés concernant l'établissement des tests, chapitre IV, section 1.2.1, étayent également cette conception). Finalement, il faut tenir compte du fait qu'il n'est pas possible, même en appliquant cette méthodologie, de saisir suffisamment toutes les facettes de la compétence de lecture. Les aptitudes suivantes en font partie.

- Non seulement comprendre les messages centraux d'un texte, mais les reconnaître comme tels (ce qui permet de distinguer l'important de l'accessoire).
- Déduire du contexte global les contenus qui ne sont pas explicités (par ex. l'intention de l'auteur).
- Saisir le sens global et l'argumentation d'un texte même si la dimension langagière pose de hautes exigences formelles et de contenu au lecteur (par ex. cheminements inédits de la pensée, complexité thématique, style inhabituel, accumulation de mots étrangers, grammaires complexe).
- Reconnaître et distinguer les contenus ou les arguments concurrents.
- Déterminer la signification d'un concept inconnu à partir du contexte textuel et/ou à l'aide de bonnes connaissances de grec et de latin.
- Comprendre un texte progressif même s'il «produit» quantité de connaissances de départ, c'est-à-dire s'il introduit continuellement de nouveaux concepts spécialisés.
- Evaluer un texte d'un œil critique (par ex. quant à ses contenus et sa pertinence relativement à une problématique spécifique).
- Saisir le signifiant d'un concept dans un contexte spécifique ou attribuer des connaissances factuelles isolées à des relations d'ordre général. Pour les informations de détail ou les notions individuelles, en particulier, il ne s'agit généralement pas tant de connaître précisément un concept que de pouvoir élaborer le contexte qui s'y rapporte.
- S'appropriier intellectuellement des textes exigeants de manière systématique, constituer la panoplie terminologique correspondante et comprendre les lignes argumentatives et les positions théoriques principales (compétences de lecture systématiques de base).
- Elaborer de nouvelles stratégies, gérer judicieusement les nouvelles connaissances, car nombre de manuels produisent un important volume de connaissances de départ (par ex. les manuels obligatoires en sociologie ou les documents didactiques publiés sur la plateforme d'apprentissage basée sur internet).

Cependant, les connaissances factuelles spécifiques aux concepts que les textes didactiques sélectionnés comprenaient ont été relativement bien saisies. Souvent, le codage large a permis d'identifier les domaines thématiques généraux afférents aux concepts. Lors de la conception des grilles de compétences du module B et de la construction des tâches du module C, on a dû tenir compte du fait que, si ces connaissances notionnelles sont indispensables pour traiter les textes didactiques, elles n'y suffisent souvent pas ou les compétences de lecture décrites ci-dessus, qui permettent de les compléter, y sont également nécessaires.

2.2 Retour des données dans les sous-projets A1 et A3

Les documents reçus comprennent l'ensemble des 16 disciplines choisies, bien que tous les enseignants contactés n'aient pas mis leurs supports d'enseignement et d'examens à disposition (malgré plusieurs rappels dans certains cas). Au total, on a pu examiner les supports de 26 disciplines d'études sur 33 filières initialement prévues. Le volume des supports analysés correspond ainsi, en se référant à l'année académique 2004/2005, à environ 31% de l'effectif total des étudiants de la Suisse (part visée: 37%)

S'agissant de l'analyse des premiers examens dans le cadre du SP A3, il faut encore noter la restriction suivante: chaque discipline ne présente pas des examens, car dans nombre de filières, en raison de la structure des études, aucun examen n'avait lieu au cours des deux premiers semestres soumis à l'analyse.

2.3 Evaluations dans les sous-projets A1 et A3

La démarche consistait donc à analyser les supports d'enseignement des premier et deuxième semestres d'études de la filière considérée et à documenter les unités de sens ou de connaissances (généralement par chapitre ou par tâche et examen) dans les grilles décrites. Plus de 1000 grilles ont ainsi été établies au total. Les éléments de connaissance et de capacité identifiés ont été portés dans une vaste banque de données sur la base des tableaux de codage. Cette banque de données permet, pour les unités de sens attribuées aux diverses disciplines de maturité, des regroupements et des évaluations tant sur le plan des divers enseignements analysés, de l'ensemble de la discipline d'études qu'au niveau de toutes les disciplines d'études analysées.

En outre, les grilles ont fait apparaître une relation entre les disciplines d'études aux universités et les branches scolaires dans les gymnases, qui permet d'observer l'interface entre ces deux institutions dans le système suisse: l'ensemble des connaissances acquises dans toutes les branches scolaires doit ainsi qualifier pour l'étude de toutes les disciplines dans les universités suisses. Le regroupement des grilles dans la banque de données a permis de relier l'un à l'autre les deux niveaux des disciplines d'études aux universités et les branches d'enseignement dans les gymnases (branches de maturité). Ces deux niveaux ne sont pas reliés symétriquement et linéairement et présentent un tissu complexe de connexions. La présentation sous forme de matrice devrait avoir permis d'échapper au danger de confondre les niveaux ou de substituer les disciplines d'études et les branches d'enseignement (par ex. comme prolongement ininterrompu d'une formation spécifique à une branche dans la discipline correspondante). Outre son éventuelle fonction propédeutique pour une discipline d'études donnée², une branche d'enseignement particulière sert avant tout à préparer à de nombreuses disciplines d'études, respectivement à toutes les disciplines d'études possibles.

Après avoir réuni, dans les 16 disciplines visées, les connaissances et capacités pertinentes pour les études, il était possible de traiter, branche d'enseignement par branche d'enseignement, les contenus préparant aux études pour l'ensemble des disciplines d'études, afin de les intégrer dans une grille de compétences cohérentes pour chaque branche d'enseignement. Il en résulte, pour chaque branche d'enseignement, une description de la contribution qu'elle doit apporter à l'acquisition de la compétence d'études (toujours dans le cadre des supports effectivement examinés). Cette description ne représente pas simplement le plan d'études cadre et les plans d'études des écoles de maturité. La figure II.1 ci-après met en évidence le lien, tel qu'il résulte des projets A et B, entre les branches scolaires et les disciplines d'études. Les disciplines d'études sont en abscisse et les branches d'enseignement se trouvent en ordonnée.

² En fait, l'art. 5 du RRM de 1995 ne prévoit pas cette fonction (pas de spécialisation), mais elle est pourtant admise par les représentants de certaines disciplines d'études.

Figure II.1: Liens entre les disciplines d'études et les branches d'enseignement

Connaissances et capacités en...	→ Mathématiques	Langue première	Géographie	Physique	Biologie	Discipline Y
... pour la discipline d'études						
<i>Germanistique</i>	GM	GLI	GG	GP	GB	GY
<i>Droit</i>	DM	DLI	DG	DP	DB	DY
<i>Sociologie</i>	SM	SLI	SG	SP	SB	SY
<i>Biologie</i>	BM	BLI	BG	BP	BB	BY
X	XM	XLI	XG	XP	XB	XY
	↓	↓	↓	↓	↓	↓
Compétences d'études à développer en...	Mathématiques	Langue première	Géographie	Physique	Biologie	Discipline Y ou compétences transversales
	↑	↑			↑	
Modèles de compétences élaborés pour les mathématiques, la langue première et la biologie						

Légende: GM = connaissances et capacités à acquérir au gymnase en mathématiques (M) pour étudier la germanistique (G) à l'université, etc.

2.4 Résultats de l'évaluation des sous-projets A1 et A3; interprétation

En raison de l'ampleur de la tâche, il n'est pas possible en cet endroit de reporter dans les tableaux de codage et dans la banque de données toutes les connaissances et capacités relevées (par branches d'enseignement).

Globalement, pour l'ensemble des disciplines d'études, les analyses qualitatives permettent de conclure aux tendances suivantes quant aux connaissances et capacités issues des diverses branches de maturité:

- Sur le plan de la catégorisation cognitive, on trouve au total nettement plus d'éléments du type connaissances factuelles (AA) que de connaissances procédurales ou métacognitives, qui sont assez rarement supposées acquises. Cette fréquence s'observe en particulier pour les éléments relevant des langues enseignées à l'école et concerne avant tout les trois langues premières (allemand, français, italien) et l'anglais. Beaucoup d'entre eux sont toutefois également donnés par le latin et le grec, une grande partie relevant simultanément de la langue première. Cette relation reflète celle des langues anciennes avec la maîtrise des concepts abstraits dans nombre de branches scientifiques. Cependant, il est aussi possible de comprendre ou de connaître ces concepts sans avoir étudié les langues anciennes au gymnase.
- S'agissant des disciplines scolaires, la catégorisation est très variable et, comme on pouvait s'y attendre, dépend étroitement de la branche analysée. Dans l'analyse de la biologie comme discipline universitaire, par exemple, on trouve fréquemment des éléments relevant de la biologie et de la chimie, enseignées à l'école; ou encore, en sciences économiques, on constate que l'enseignement universitaire présuppose un éventail de notions économiques déjà étonnamment étendu. Ainsi, par exemple, l'étudiant qui a choisi l'option spécifique biologie/chimie ou économie/droit aura certains avantages lorsqu'il entamera ses études de biologie ou d'économie.
- Les connaissances de grec et de latin peuvent le plus souvent passer pour utiles, mais ne sont pas nécessairement considérées comme indispensables, par exemple pour les études en droit, en histoire ou en pédagogie; une connaissance approfondie du latin et du grec pourrait néanmoins, on l'a vu, présenter généralement des avantages surtout pour le vocabulaire, qui est lui-même très important pour l'ensemble des branches d'études. Mais il faut tenir compte du fait que les termes techniques peuvent être compris ou acquis même sans études linguistiques de ce type.

- Nous devons en l'occurrence laisser ouverte la question de savoir si les remarques précédentes plaident en faveur d'une réintroduction obligatoire des langues anciennes ou si la voie directe de l'appropriation du vocabulaire spécialisé conduit à une meilleure aptitude aux études.
- Outre les connaissances importantes de la langue première et un riche vocabulaire étranger, on exige dans certains cas des étudiants une bonne compréhension de l'écrit: notamment en sciences humaines et sociales, en histoire, en sociologie et en pédagogie, on observe une concentration d'éléments concernant le traitement de textes complexes et requérant des compétences adéquates en lecture.
- S'agissant des connaissances en sciences expérimentales, on ne constate pratiquement aucune différence quant au nombre d'éléments de physique et de chimie requis par chacune des branches non expérimentales (c'est-à-dire, dans notre cas, toutes les branches sauf la physique, la biologie et la médecine). En revanche, des notions de biologie sont nettement plus souvent exigées dans les sciences humaines et sociales que des concepts de physique ou de chimie.
- Des acquis sont également attendus dans le domaine de compétence interdisciplinaire que constitue le travail scientifique. Mais des cours ou séminaires d'introduction au travail scientifique sont également proposés dans certaines branches.
- L'analyse des connaissances requises dans le cadre des exercices et examens montre que, le plus souvent, ces derniers font peu référence à des éléments supposés acquis dès le gymnase; parmi les compétences cognitives requises lors des examens, on trouve surtout des types de compétences comme «mémoriser les connaissances», «comprendre», «appliquer» et «analyser». Par contre, on exige assez peu des étudiants qu'ils sachent «évaluer, créer» ou «générer, planifier», du moins pendant les deux premiers semestres d'études.

Les résultats obtenus dans les sous-projets A1 et A3 devaient essentiellement fournir une base précieuse aux tests réalisés dans la suite du projet pour la langue première, les mathématiques et la biologie enseignées à l'école (cf. modules B et C, respectivement chapitres III et IV du présent rapport).

3 Sous-projet A4: enquête auprès des enseignants

3.1 Le questionnaire

Selon le projet de mandat (cf. Eberle, 2005), dix entretiens dirigés étaient prévus, chacun avec deux professeurs ou conseillers aux études des facultés des Lettres, des Sciences, de Médecine, d'Economie et de Droit. Pendant la phase de collecte des supports, dans le cadre du projet A1, il est cependant apparu qu'il serait difficile de trouver les personnes adéquates. Celles-ci auraient dû répondre au profil suivant:

- Elles devraient avoir la vue d'ensemble des contenus des disciplines de tous les enseignements des deux premiers semestres.
- Pour éviter des lieux communs aux questions portant sur les connaissances et les capacités de départ nécessaires pour les études, ces personnes devraient avoir étudié la didactique. Autrement dit, elles devraient pouvoir clarifier les questions suivantes lors de la planification des enseignements: de quels prérequis (connaissances et capacités) les étudiants doivent-ils disposer dans les divers enseignements? Comment peut-on évaluer l'état des connaissances et des capacités des étudiants?

Dans le cadre de la collecte des documents pour les autres sous-projets, il est apparu que seules très peu de personnes disposaient des connaissances nécessaires aux entretiens pour tous les enseignements de la discipline d'études correspondante. C'est pourquoi, d'entente avec le groupe de pilotage, on a pris la décision d'impliquer dans l'enquête tous les enseignants des enseignements étudiés dans le cadre des sous-projets A1 et A3. Apprécier les connaissances et les capacités requises pour son propre enseignement devrait être dans la plupart des cas plus aisé pour les intéressés eux-mêmes. Comme le nombre potentiel des personnes

interviewées passait de ce fait à plus de 300, l'enquête a dû se faire par écrit. Par rapport à une interview orale, l'écrit présente certes des inconvénients, mais il a aussi des avantages. Les principaux avantages résident dans le fait que l'on peut mieux standardiser l'enquête et que les enseignants interrogés disposent de plus de temps de réflexion.

Le but de l'enquête était d'apprendre quelles connaissances et capacités spécifiques et interdisciplinaires les enseignants supposent acquises chez leurs étudiants des premier et deuxième semestres. Les enseignants étaient également invités à mentionner quelles lacunes spécifiques ou générales ils observaient chez les étudiants débutants.

3.2 Base de données du SP A4 et déroulement de l'enquête

L'enquête a été conduite dans des universités de Suisse alémanique avec les enseignants responsables des enseignements des deux premiers semestres dans les disciplines déterminées dans les SP A1 et A3. La limitation à la Suisse alémanique, définie d'entente avec le groupe de pilotage, découlait de considérations économiques concernant les ressources. Le relevé a été complété par une enquête conduite auprès d'enseignants universitaires bilingues sélectionnés au Tessin et en Suisse romande, lesquels ont estimé par la suite que les résultats seraient en majeure partie transposables à leur région linguistique. Par ailleurs, on peut noter rétrospectivement que les résultats obtenus de l'enquête auprès des enseignants n'ont influencé que de manière insignifiante la construction du test du module C: ils devraient seulement servir en l'occurrence d'éventuelle correction approximative, car il devrait être difficile pour nombre d'enseignants de déterminer précisément les connaissances et capacité de départ requises par leurs enseignements. Une telle correction approximative n'a pas été nécessaire, à l'exception du relevé complémentaire du domaine de compétences «fonctionnement de la langue / compétence grammaticale et orthographe». Cependant, les résultats de l'enquête sont intéressants comme résultats autonomes en vue de réfléchir au développement futur du gymnase et ils sont des aides utiles à l'interprétation dans les sous-projets C et D1. Des analyses ultérieures permettront peut-être encore de procéder à des comparaisons transversales entre les données fournies par les enseignants dans le questionnaire et les unités de sens saisies dans la base de données suite au codage des supports didactiques de ces enseignants.

Au total, 310 questionnaires ont été adressés aux enseignants responsables d'enseignements au premier ou au deuxième semestre et qui sont actifs dans le cadre des filières d'études sélectionnées, mentionnées au tableau II.1, et les universités choisies pour le SP A. 140 questionnaires ont été remplis, ce qui représente un taux de retour de 45 %. Les enseignants se répartissent comme suit en fonction des disciplines d'études: germanistique (20), anglais (9), histoire (16), psychologie (7), pédagogie (12), sociologie (3), sciences politiques (4) communication & médias (5), économie (10), droit (16), informatique (5), physique (5), biologie (16), médecine (5), architecture (7).

3.3 Résultats de l'enquête du SP A4

Voici une présentation des principaux résultats. Les réponses vides ne sont pas mentionnées spécialement.

3.3.1 Signification des disciplines gymnasiales pour les enseignements (question 1)

On a tout d'abord demandé aux enseignants d'évaluer quelle importance ils attribuaient aux connaissances et capacités spécifiques aux différentes disciplines de maturité gymnasiale et aux domaines de l'utilisation de l'informatique et de la science informatique pour comprendre et suivre avec succès leur enseignement. La question suivante leur a été posée:

«Dans quelles branches gymnasiales les étudiants doivent-ils déjà posséder des connaissances et des capacités spécifiques pour être en mesure de suivre votre enseignement? Veuillez vous référer à cet égard aux connaissances et capacités spécifiques à une discipline dont les étudiants doivent disposer dès le commencement de leurs études et

qu'ils ne peuvent donc pas acquérir dans d'autres enseignements de la filière. Veuillez marquer d'une croix le champ le plus pertinent par discipline.»

Les résultats se trouvent dans le tableau II.5. Outre les pourcentages selon les réponses possibles, la dernière colonne indique pour chaque branche la moyenne constituée à partir de l'échelle à quatre niveaux. Les lignes sont classées en fonction de cette valeur.

L'importance attribuée à la formation préalable pour le propre enseignement varie beaucoup selon les disciplines. La langue première, l'anglais, l'utilisation de l'informatique et les mathématiques (dans cet ordre) reçoivent les valeurs les plus élevées quant à l'importance. La biologie et la physique se trouvent au milieu de ce classement et la chimie arrive au début du dernier tiers. Ainsi, l'enquête conforte la décision prise de retenir la langue première et les mathématiques dans la sélection des disciplines testées et la biologie comme discipline scientifique.

Tableau II.5: Importance des branches gymnasiales pour le propre enseignement

Branches gymnasiales ou domaines	Aucune connaissance ni capacité (1)	Connaissances et capacités fragmentaires (2)	Connaissance des principales bases (3)	Connaissances et capacités approfondies (4)	Moyenne
Langue première	5.2%	6.0%	25.4%	63.4%	3.47
Anglais	16.7%	11.9%	43.7%	27.8%	2.83
Utilisation de l'informatique	19.7%	18.9%	45.9%	15.6%	2.57
Mathématiques	36.1%	15.6%	32.8%	15.6%	2.28
Histoire	36.0%	22.4%	28.0%	13.6%	2.19
Philosophie	41.8%	30.3%	27.9%	0.0%	1.86
Français	54.2%	15.8%	22.5%	7.5%	1.83
Economie & droit	50.8%	26.7%	19.2%	3.3%	1.75
Géographie	56.0%	27.6%	14.7%	1.7%	1.62
Biologie	68.1%	11.8%	13.4%	6.7%	1.59
Physique	70.1%	10.3%	14.5%	5.1%	1.55
Ethique & religion	60.8%	30.8%	8.3%	0.0%	1.47
Latin	64.5%	28.1%	7.4%	0.0%	1.43
Science informatique	72.0%	18.6%	7.6%	1.7%	1.39
Psychologie	72.5%	18.3%	8.3%	0.8%	1.38
Chimie	78.8%	12.7%	2.5%	5.9%	1.36
Italien	77.8%	13.7%	8.5%	0.0%	1.31
Pédagogie	81.2%	12.0%	5.1%	1.7%	1.27
Arts visuels	83.8%	11.1%	4.3%	0.9%	1.22
Grec	89.5%	9.6%	0.9%	0.0%	1.11
Musique	91.4%	8.6%	0.0%	0.0%	1.09

N max. = 140

Si l'on ventile les réponses à la question selon les diverses disciplines d'études, les positions se modifient en partie (cf. tableau II.6). Ci-après sont mentionnées chaque fois les quatre premières positions et celles de disciplines testées dans EVAMAR (avant que la discipline scientifique ne soit déterminée).

- Pour la germanistique, les branches désignées comme les plus importantes sont la langue première, l'utilisation de l'informatique, l'histoire et l'anglais. La biologie, la physique et la chimie se classent en 10^e, 12^e et 18^e position, tandis que les mathématiques sont au 15^e rang.

- S'agissant de l'anglais, les quatre branches jugées les plus importantes sont l'anglais, la langue première, l'utilisation de l'informatique et l'histoire. La biologie et la physique occupent la 11^e place, tandis que les mathématiques sont au 17^e rang. Aucune connaissance ni capacité n'est attendue en chimie.
- Les enseignants d'histoire ont jugé que la langue première, l'histoire, l'anglais et le français sont les branches les plus importantes pour étudier l'histoire. Les mathématiques et la biologie arrivent en 12^e et en 17^e position. Des connaissances et capacités ne sont pas attendues en chimie et en physique.

Tableau II.6: Importance des branches gymnasiales en fonction des disciplines d'études (classement par ordre d'importance entre parenthèses)

Branches gymnasiales ou domaines	Germanis- tique	Anglais	Histoire	Psycholo- gie	Pédagogie	Sociologie	Sciences politiques	Communi- cation & médias	Economie	Droit	Infor- matique	Physique	Biologie	Médecine	Architec- ture
Langue première	3.60 (1)	2.78 (2)	3.88 (1)	3.83 (1)	3.45 (1)	2.67 (4)	3.75 (1)	3.75 (1)	3.50 (1)	3.93 (1)	2.80 (2)	2.25 (3)	3.47 (1)	3.75 (1)	2.67 (1)
Mathématiques	1.23 (15)	1.14 (17)	1.44 (12)	3.17 (3)	2.00 (7)	3.33 (2)	2.75 (5)	3.00 (2)	2.64 (2)	1.85 (9)	3.60 (1)	3.40 (1)	3.13 (3)	2.75 (3)	2.20 (7)
Anglais	2.47 (4)	4.00 (1)	3.31 (3)	3.50 (2)	2.64 (3)	3.33 (2)	3.50 (2)	2.75 (4)	2.27 (5)	2.15 (7)	2.00 (4)	2.25 (3)	3.06 (4)	2.75 (3)	2.60 (2)
Biologie	1.38 (10)	1.29 (11)	1.06 (17)	2.00 (6)	1.40 (11)	1.00 (21)	1.25 (12)	1.00 (21)	1.10 (16)	1.08 (21)	1.20 (9)	1.25 (7)	3.24 (2)	3.25 (2)	1.00 (21)
Chimie	1.15 (18)	1.00 (21)	1.00 (21)	1.33 (11)	1.00 (21)	1.00 (21)	1.00 (21)	1.00 (21)	1.10 (16)	1.08 (21)	1.40 (6)	1.25 (7)	2.63 (5)	2.75 (3)	1.00 (21)
Physique	1.31 (12)	1.29 (11)	1.00 (21)	1.50 (10)	1.00 (21)	1.00 (21)	1.00 (21)	1.00 (21)	1.20 (12)	1.15 (14)	2.20 (3)	2.80 (2)	2.56 (6)	2.75 (3)	2.00 (8)
Utilisation de l'informatique	2.50 (2)	2.63 (3)	2.75 (5)	2.50 (4)	3.00 (2)	3.67 (1)	3.00 (4)	3.00 (2)	2.40 (4)	2.43 (5)	1.40 (6)	2.00 (5)	2.56 (6)	2.33 (7)	2.60 (2)
Psychologie	1.36 (11)	1.29 (11)	1.06 (17)	2.17 (5)	2.09 (6)	1.33 (9)	1.00 (21)	1.67 (8)	1.30 (11)	1.38 (12)	1.00 (21)	1.00 (21)	1.06 (18)	2.00 (8)	1.50 (11)
Philosophie	2.19 (6)	1.67 (6)	2.31 (7)	1.67 (7)	2.00 (7)	1.67 (7)	2.75 (5)	1.00 (21)	1.60 (7)	2.31 (6)	1.20 (9)	1.25 (7)	1.25 (12)	1.67 (11)	2.25 (6)
Science informatique	1.08 (21)	1.29 (11)	1.18 (15)	1.67 (7)	1.10 (14)	2.00 (5)	1.25 (12)	1.33 (9)	1.50 (8)	1.15 (14)	1.60 (5)	1.75 (6)	1.94 (8)	1.67 (11)	1.00 (21)
Pédagogie	1.15 (18)	1.14 (17)	1.19 (14)	1.60 (9)	2.45 (4)	1.33 (9)	1.00 (21)	1.00 (21)	1.20 (12)	1.08 (21)	1.00 (21)	1.00 (21)	1.06 (18)	1.33 (21)	1.25 (12)
Histoire	2.50 (2)	2.11 (4)	3.35 (2)	1.33 (11)	2.45 (4)	2.00 (5)	3.25 (3)	2.25 (5)	1.70 (6)	2.62 (3)	1.00 (21)	1.25 (7)	1.13 (16)	2.00 (8)	2.50 (4)
Economie et droit	1.23 (15)	1.14 (17)	2.25 (8)	1.17 (13)	1.60 (10)	1.67 (7)	2.25 (9)	2.25 (5)	2.45 (3)	2.50 (4)	1.40 (6)	1.25 (7)	1.19 (13)	1.33 (21)	1.75 (9)
Ethique & religion	1.46 (9)	1.44 (9)	2.00 (9)	1.17 (13)	1.78 (9)	1.33 (9)	1.50 (10)	1.00 (21)	1.50 (8)	1.69 (10)	1.00 (21)	1.00 (21)	1.06 (18)	1.75 (10)	1.25 (12)
Arts visuels	1.08 (21)	1.57 (8)	1.00 (21)	1.00 (21)	1.00 (21)	1.00 (21)	1.00 (21)	1.00 (21)	1.20 (12)	1.08 (21)	1.20 (9)	1.00 (21)	1.50 (9)	1.67 (11)	2.40 (5)
Géographie	1.62 (8)	1.29 (11)	2.69 (6)	1.00 (21)	1.22 (13)	1.33 (9)	2.50 (7)	2.00 (7)	1.40 (10)	1.46 (11)	1.20 (9)	1.00 (21)	1.44 (10)	1.67 (11)	1.75 (9)
Français	2.21 (5)	1.63 (7)	2.88 (4)	1.00 (21)	1.30 (12)	1.33 (9)	2.50 (7)	1.00 (21)	1.10 (16)	3.00 (2)	1.00 (21)	1.00 (21)	1.44 (10)	1.33 (21)	1.25 (12)
Latin	1.93 (7)	1.75 (5)	2.00 (9)	1.00 (21)	1.00 (21)	1.33 (9)	1.50 (10)	1.00 (21)	1.20 (12)	1.31 (13)	1.20 (9)	1.00 (21)	1.19 (13)	1.67 (11)	1.00 (21)
Grec	1.23 (15)	1.00 (21)	1.20 (13)	1.00 (21)	1.00 (21)	1.00 (21)	1.25 (12)	1.00 (21)	1.10 (16)	1.08 (21)	1.20 (9)	1.00 (21)	1.07 (17)	1.67 (11)	1.00 (21)
Italien	1.31 (12)	1.38 (10)	1.73 (11)	1.00 (21)	1.10 (14)	1.33 (9)	1.00 (21)	1.00 (21)	1.00 (21)	1.92 (8)	1.00 (21)	1.00 (21)	1.19 (13)	1.33 (21)	1.00 (21)
Musique	1.25 (14)	1.29 (11)	1.13 (16)	1.00 (21)	1.00 (21)	1.00 (21)	1.25 (12)	1.00 (21)	1.00 (21)	1.08 (21)	1.00 (21)	1.00 (21)	1.00 (21)	1.33 (21)	1.00 (21)

- Selon les indications des enseignants de psychologie, les études dans leur discipline requièrent principalement des connaissances et des capacités relevant des branches suivantes: langue première, anglais et mathématiques. Les connaissances en utilisation de l'informatique arrivent en quatrième position, tandis que la biologie, la physique et la chimie occupent respectivement les 6^e, 10^e et 11^e rangs.
- La langue première, l'utilisation de l'informatique, l'anglais, la pédagogie et l'histoire sont les branches jugées les plus importantes comme base d'études en pédagogie. Elles sont suivies par les mathématiques et la biologie en 7^e et 11^e position. Selon les enseignants en pédagogie, des connaissances et capacités en chimie et en physique ne sont pas requises.
- Les enseignants en sociologie placent au premier plan l'utilisation de l'informatique, les mathématiques, l'anglais et la langue première, tandis qu'ils jugent la physique, la chimie et la biologie sans importance.
- Des études en sciences politiques requièrent avant tout, selon les enseignants de cette discipline, des connaissances et des capacités en langue première, en anglais, en histoire et en utilisation de l'informatique. On trouve les mathématiques au 5^e rang et la biologie en 12^e position. Des connaissances de physique et de chimie ne sont pas nécessaires.
- Les enseignants en communication & médias placent au plus haut degré d'importance de vastes connaissances et capacités en langue première, mathématiques, anglais et utilisation de l'informatique. La biologie, la chimie et la physique n'ont pas d'importance selon eux.
- Les enseignants en économie jugent particulièrement importantes, pour les études dans leur discipline, les connaissances et capacités en langue première, mathématiques, économie & droit et utilisation de l'informatique. La physique (12^e rang) est jugée moins importante, tandis que la chimie et la biologie (16^e rang l'une et l'autre) ne revêtent presque aucune importance.
- En droit, les branches que sont la langue première, le français, l'histoire et économie & droit occupent les premières places. En revanche, les enseignants en droit estiment qu'il ne faut guère de connaissances de départ issues de la physique (14^e rang), de la chimie ou de la biologie (21^e rang chacune).
- Des études d'informatique requièrent, selon les enseignants concernés, avant tout des connaissances et des capacités en mathématiques, en langue première, en physique et en anglais. Suivent la chimie (6^e rang) et la biologie (9^e rang).
- Selon les enseignants de physique, les études dans leur discipline font surtout appel aux connaissances et capacités issues des mathématiques, de la physique, de l'anglais et de la langue première. La chimie et la biologie sont jugées d'importance égale (7^e position).
- En biologie, les enseignants attribuent la plus grande importance aux connaissances et capacités dans les branches suivantes: langue première, biologie, mathématiques et anglais, immédiatement suivies par la chimie et la physique (5^e et 6^e rangs).
- Les enseignants de médecine jugent que les branches scolaires suivantes sont les plus importantes dans leur enseignement: langue première (1^{er} rang), biologie (2^e rang), anglais, physique, chimie et mathématiques (toutes au 3^e rang).
- Enfin, en architecture, la langue première, l'utilisation de l'informatique, l'anglais et l'histoire occupent les quatre premières positions, tandis que les mathématiques et la physique arrivent au 7^e et au 8^e rang. La chimie et la biologie sont jugées sans importance.

Globalement, la langue première demeure la branche gymnasiale jugée majoritairement la plus importante si l'on compare les moyennes des estimations sur l'ensemble des disciplines d'études. L'anglais et l'utilisation de l'informatique sont aux premiers rangs. Pour les autres branches, on note des différences de classement considérables dans certains cas. Dans les disciplines linguistiques et en sciences humaines et sociales, les mathématiques sont jugées nettement moins importantes, comme on s'y attendait, tout en revêtant une forte importance dans de nombreuses disciplines d'études. On relève également en tête de classement la branche de maturité correspondant directement à la discipline d'études elle-même.

3.3.2 Contenus concrets des connaissances et capacités dans les branches gymnasiales importantes pour l'enseignement visé (question 2)

Les enseignants ont ensuite eu à répondre à la question ouverte suivante concernant les connaissances et capacités concrètes sur lesquelles repose leur enseignement:

«Veuillez décrire aussi précisément que possible, pour chaque branche désignée dans votre réponse à la question 1 comme requérant des «connaissances ou capacités approfondies» ou des «connaissances des principales bases», quelles connaissances et capacités vous supposez acquises dans votre enseignement. Veuillez ne décrire également que les connaissances et capacités dont les étudiants doivent disposer dès le commencement de leurs études et qu'ils ne peuvent donc pas acquérir dans d'autres enseignements de l'université/EPF.»

Le tableau II.7 présente une récapitulation synthétique des résultats. Les chiffres indiquent le nombre de mentions équivalentes³.

Tableau II.7: Connaissances et capacités spécifiques requises: mentions expresses des attentes par discipline

Branches	Contenus
Langue première	<ul style="list-style-type: none"> • Maîtrise de la grammaire: 7 enseignants de germanistique; 7 autres enseignants, chacun d'une discipline d'études différente. • Maîtrise de la langue: 55 mentions parmi l'ensemble des disciplines d'études, dont près de 100 % des enseignants de droit; très bonne maîtrise de la langue notamment du point de vue stylistique: 20 enseignants de diverses disciplines; formulation scientifique: 2 enseignants de biologie. • Syntaxe (6); orthographe (8); ponctuation (3); sémantique (2). • Selon les enseignants en germanistique: connaissance élémentaire de l'histoire de la langue (1), aperçu des principales époques de l'histoire littéraire (6), connaissance de quelques méthodes fondamentales d'analyse des textes et d'interprétation (6), lecture des textes d'une liste de référence («canon littéraire») (3), connaissance des notions de base de la recherche littéraire (2). • Notions essentielles de l'analyse du discours narratif, de la théorie du théâtre, de la poésie et de l'histoire de la littérature: 2 enseignants d'anglais; bases de l'interprétation des énoncés littéraires: 1 enseignant d'anglais. • Mentionné en outre par des enseignants en germanistique: bases de la linguistique (1); expérience générale de textes exigeants ou littéraires (3); bases de la structure linguistique de la langue première et concepts grammaticaux de base (2); connaissance des types de textes, des formes de médias et de communication (1); connaissance des principales règles de la poétique occidentale (1); modèles fondamentaux de la théorie de la lecture (connaissance passive et active) (1); connaissances textuelles (1) et précision de la lecture (1), complété par 1 mention venue de la discipline du droit.
Mathématiques	<ul style="list-style-type: none"> • Calcul infinitésimal: 5 enseignants en économie, 3 en informatique, 7 en biologie, 3 en médecine et 6 enseignants de diverses autres disciplines; calcul différentiel seulement: 4 enseignants de diverses filières; calcul intégral seulement: 1 enseignant des sciences de la communication. • Connaissances approfondies de l'algèbre: 3 enseignants en économie, 11 enseignants d'autres filières; fonctions exponentielles et logarithmiques: 4 enseignants en biologie; calcul des probabilités: 4 enseignants en psychologie, 7 enseignants d'autres filières; statistiques et leur interprétation: 4 enseignants en biologie, 7 enseignants d'autres filières. • Théorie des ensembles (3); logique (6); planimétrie (2); géométrie vectorielle (7); trigonométrie (6); nombres complexes (2); théorie des fonctions (7); arithmétique (6, dont 3 en informatique); systèmes de numération (2). • Connaissances élémentaires de la théorie des jeux (1); suites et séries (2); polynômes (1); théorie combinatoire (2); distribution binomiale (1); équations différentielles (1); notions de preuve (1); fonctions d'onde (1).

³ Si aucune indication supplémentaire n'accompagne le chiffre, les mentions proviennent d'enseignants de diverses disciplines.

Branche	Contenus
Biologie	<ul style="list-style-type: none"> Concepts fondamentaux de l'évolution (1); développement anatomique ou physiologique de l'être humain: 2 enseignants issus chacun de filières autres que les sciences expérimentales. Orientation des études vers la médecine: génétique, génétique moléculaire, développement des cellules végétales et animales, cycles des substances, écologie, éléments de microbiologie et de théorie de l'évolution: 1 mention; liens fondamentaux de la biologie moléculaire de la génétique moléculaire et de la biochimie: 1 mention. Mentions issues d'enseignants en biologie: fondements de la division cellulaire (2); photosynthèse (2); botanique (1); zoologie (1); structure de la cellule (1); concepts de base de l'écologie (4); biologie de l'évolution (2); génétique de l'évolution comme principe (1); génétique classique (4); génétique moléculaire (4); biologie moléculaire (2); neurobiologie (1); physiologie (3); connaissances des organismes (1); crise de la biodiversité (1); concepts d'espèces (1); métabolisme (1); respiration (1); nutrition (1); reproduction (1); membranes (1); biologie cellulaire (2); évolution (5); biochimie des protéines (1); bases de l'anatomie (1).
Grec	<ul style="list-style-type: none"> Connaissance de base: 3 enseignants en germanistique et en histoire. Maîtrise de l'alphabet grec: 1 enseignant en communication.
Latin	<ul style="list-style-type: none"> Connaissances de base (5); connaissances visant la préparation d'analyse linguistiques structurelles (1); vocabulaire élémentaire (2). Connaissances approfondies permettant de traiter les sources originales: 2 enseignants en histoire.
Français	<ul style="list-style-type: none"> Maîtrise de la langue: 4 enseignants en histoire, 1 enseignant en droit; connaissances permettant de traiter des textes spécifiques au domaine: 7 enseignants en droit, 5 en histoire et 1 en sciences politiques; compréhension de l'écrit (3).
Anglais	<ul style="list-style-type: none"> Haut niveau de maîtrise de la langue (au moins C1 dans le Portefeuille européen des langues): tous les enseignants d'anglais interrogés. Capacité de compréhension auditive, connaissance des divers niveaux de langue, notions essentielles de l'analyse du discours narratif, de la théorie du théâtre, de l'histoire de la littérature: chacune 1 enseignant (différent) en anglais. Connaissances linguistiques orales et écrites au niveau voulu pour traiter la littérature spécialisée (54).
Italien/espagnol	<ul style="list-style-type: none"> Connaissance de l'italien permettant de traiter les prononcés de jugement: 3 enseignants en droit. Connaissances passives de l'italien: 1 enseignant en histoire.
Histoire	<ul style="list-style-type: none"> Large connaissances de base: 5 enseignants en germanistique, 4 en anglais, en histoire et en droit. Aperçu général de l'histoire suisse et européenne moderne: 5 enseignants en histoire, 10 enseignants de diverses filières; connaissances approfondies du Moyen-Age: 1 enseignant en germanistique; extension au continent nord-américain: 1 enseignant d'anglais. Connaissance de l'histoire politico-sociale récente (5). Notions de base, un certain bagage d'années repères et la capacité de gérer les sources: 4 enseignants en histoire; connaissances de la période comprise entre l'Antiquité et l'époque actuelle: 1 enseignant en histoire et 1 enseignant en sciences politiques.
Economie & droit	<ul style="list-style-type: none"> Connaissances fondamentales (cycle économique, offre/demande, droit, morale): 3 enseignants en histoire, 3 enseignants en économie, 4 enseignants d'autres filières; compréhension approfondie de la structure du droit public et du droit privé: 3 enseignants en droit. Connaissances du système économique: 1 enseignant en sociologie; titres, intérêts, intérêts composés et mise en rapport des concepts de coûts et d'utilité: 1 enseignant en informatique.
Géographie	<ul style="list-style-type: none"> Orientation dans l'espace (par ex. espace méditerranéen): 3 enseignants en histoire; aspects de géographie humaine (modification des espaces vitaux dans la dimension historique, par ex. aménagement du territoire): 1 enseignant en histoire. Connaissances du climat, des zones climatiques des continents, de la météorologie et de la tectonique des plaques: 2 enseignants en biologie.
Chimie	<ul style="list-style-type: none"> Connaissances approfondies de la chimie (chimie générale, atomes, liaisons chimiques, modèle VSEPR, équilibres, tableau périodique des éléments (TPE), chimie des éléments, chimie organique, structure de l'atome, molécule, forces de liaison, sels, métaux, réactions chimiques, équations stœchiométriques, thermodynamique, acides et bases, réactions d'oxydoréduction, poids de formule, poids atomique, ion, mole, nombre d'oxydation, groupes fonctionnels (alcools, acides carboniques, cétones, aldéhydes, amines, amides): 2 enseignants en médecine et 6 enseignants en biologie.

Branche	Contenus
Physique	<ul style="list-style-type: none"> • Connaissances d'acoustique et de théorie des ondes: 1 enseignant en germanistique et 1 enseignant en anglais. • Connaissances utiles à l'estimation des ordres de grandeur, unités SI, bases en mécanique, optique, acoustique, thermodynamique et électricité: 7 enseignants en biologie, 3 enseignants en physique et 4 autres enseignants de disciplines influencées par les sciences expérimentales.
Philosophie	<ul style="list-style-type: none"> • Connaissances fondamentales: 3 enseignants en germanistique, 7 enseignants d'autres disciplines diverses; connaissances des principaux récits bibliques: 3 enseignants en histoire, 1 enseignant en pédagogie. • Ecoles de pensée spécifiques (en particulier Socrate, Platon, Descartes, Hume, Leibniz, Kant, Humboldt, Hegel, Marx, Nietzsche, Wittgenstein, Sartre): 4 mentions. • Connaissance de la philosophie et de l'histoire de la philosophie des 19^e et 20^e siècles: 1 enseignant en architecture; connaissances fondamentales des techniques argumentatives: 1 enseignant en sciences politiques. • Vue d'ensemble de l'histoire des idées: 1 enseignant en droit et 1 enseignant en architecture. • Choix de pédagogues classiques, résultats de la recherche familiale et scolaire, bases de psychologie du développement, délimitation de la pédagogie et des autres disciplines: 2 enseignants en pédagogie.
Informatique	<ul style="list-style-type: none"> • Connaissances utilisateur permettant l'utilisation de progiciels d'application usuels (programmes MS-Office): 45 mentions de presque toutes les filières. • Travail productif à l'aide de banques de données (pas seulement avec Google) et d'algorithmes (7); programmation proprement dite (4); gestion d'un système d'exploitation: 1 enseignant en informatique.
Musique	<ul style="list-style-type: none"> • Aucune mention de connaissance ou capacité de départ attendue.
Arts visuels	<ul style="list-style-type: none"> • Connaissances fondamentales de l'histoire de l'art: 1 enseignant en anglais; connaissance de l'analyse des signes visuels, images et photographies: 1 autre enseignant en anglais; capacité de représenter visuellement un organisme: 1 enseignant en biologie; capacité de dessiner à partir de sa propre conception: 1 enseignant en architecture.

Les réponses apportées à la question posée indiquent notamment que des divergences existent quant à la manière de concevoir le programme de la branche de maturité considérée. Lorsque les enseignants mentionnent la grande importance de la langue première, ils pensent pour la plupart à la maîtrise de la langue, à la grammaire, à la syntaxe, etc.. Mais seuls quelques germanistes et un angliciste ont mentionné dans leurs attentes les véritables contenus essentiels de la germanistique que sont l'histoire de la langue ou de la littérature, les méthodes linguistiques d'analyse et d'interprétation des textes, les fondements de la linguistique, la théorie du théâtre, etc. En d'autres termes, la désignation «langue première» véhicule des contenus spécifiques différents dans l'esprit des répondants. Si la langue première est considéré généralement comme la branche de maturité la plus importante, c'est donc principalement la compétence linguistique en langue première qui est visée. Il en va de même de l'anglais: l'accent est mis sur la capacité à traiter de la littérature spécialisée en langue anglaise (et non pas sur la compétence en histoire littéraire).

3.3.3 Lacunes actuellement constatées dans les connaissances et capacités (question 3)

La question ci-après a permis de demander aux enseignants quelles lacunes ils identifient parmi les étudiants dans les connaissances et capacités précédemment désignées comme importantes.

«Des connaissances et capacités spécifiques désignées à la question 2, quelles sont celles, qui permettraient de suivre votre enseignement, dont ne disposent généralement pas (ou plus) les étudiants au début de leurs études?»

Les résultats sont présentés au tableau II.8, également sous forme synthétique, avec le nombre de mentions⁴.

Tableau II.8: Connaissances et capacités spécifiques requises: mentions expresses des lacunes par discipline

Branche	Contenus
Langue première	<ul style="list-style-type: none"> Maîtrise de la langue (dans certains cas sans explication plus précise quant aux aspects de la langue insuffisamment maîtrisés): 11 enseignants en droit, 3 enseignants en histoire, 6 enseignants de filières non axées sur les langues. Grammaire: 5 enseignants en germanistique, 6 enseignants en histoire, 2 autres enseignants de filières différentes; orthographe: 4 enseignants en germanistique, 4 enseignants d'autres filières; syntaxe: 3 enseignants en histoire, 1 enseignant en germanistique; ponctuation: 4 enseignants en histoire, 1 enseignant en germanistique. Connaissance de l'histoire de la littérature: 5 enseignants en germanistique; connaissances utiles à l'interprétation de textes littéraires: 3 enseignants de diverses filières; connaissances en linguistique: 1 enseignant en germanistique et 1 enseignant en anglais. Rédaction de textes scientifiques, gestion des textes scientifiques: 1 enseignant en sciences politiques et 1 enseignant en anglais. Panoplie terminologique pour l'analyse d'énoncés littéraires: 3 enseignants en germanistique; connaissances de l'étymologie: 1 enseignant en germanistique. Elaboration de textes complexes: 1 enseignant en économie. «Analphabétisme structurel»: 1 enseignant en droit.
Mathématiques	<ul style="list-style-type: none"> Statistiques, calcul des probabilités (7). Calcul infinitésimal (5); calcul intégral (4). Nombres complexes: 1 enseignant en informatique et 1 enseignant en physique; géométrie vectorielle: 4 enseignants de filières différentes. Fonctions exponentielles et logarithmiques (3). Algèbre: 1 enseignant en économie et 1 enseignant en informatique. Logique propositionnelle (3); preuves: 1 enseignant en informatique. Arithmétique: 1 enseignant en informatique et 1 enseignant en médecine. Trigonométrie: 1 enseignant en architecture. Théorie des fonctions: 1 enseignant en économie et 1 enseignant en physique. Système de numération binaire: 1 enseignant en physique. Traitement des suites et des séries: 1 enseignant en physique et 1 enseignant en architecture. Fonctions trigonométriques: 1 enseignant en médecine. Equations quadratiques et théorème de Thalès: 1 enseignant en architecture. Théorie des nombres et géométrie analytique: 1 enseignant en informatique.
Biologie	<ul style="list-style-type: none"> Application de connaissances d'anatomie à la linguistique: 1 enseignant en anglais. Connaissances de base concernant le système nerveux végétatif, la cognition, la neuro-anatomie: 1 enseignant en pédagogie.
Grec	<ul style="list-style-type: none"> Connaissances (de base): 1 enseignant en germanistique et 1 enseignant en histoire.
Latin	<ul style="list-style-type: none"> Connaissances de base (5).

⁴ Si aucune indication supplémentaire n'accompagne le chiffre, les mentions proviennent d'enseignants de diverses disciplines.

Branche	Contenus
Français	<ul style="list-style-type: none"> Connaissances permettant de traiter la littérature spécialisée: 4 enseignants en histoire, 3 enseignants en droit, 1 enseignant en sciences politiques; connaissances du vocabulaire de base spécifique à la discipline: 1 enseignant en droit.
Anglais	<ul style="list-style-type: none"> Niveau de maîtrise de la langue et compréhension des nuances verbales: 6 enseignants en anglais; connaissances de la grammaire et vocabulaire: 1 enseignant en anglais; maîtrise de la langue sous l'angle du traitement de textes scientifiques: 2 enseignants en anglais; connaissances linguistiques: 1 enseignant en anglais. Niveau du «advanced level» à l'entrée: 1 enseignant en anglais. Exercice du traitement de textes complexes de la littérature spécialisée, vocabulaire et compréhension différenciée de la langue: 5 enseignants en pédagogie, 4 enseignants en histoire, 8 autres enseignants de filières d'études d'orientations diverses.
Italien/espagnol	<ul style="list-style-type: none"> Connaissances de base (3).
Histoire	<ul style="list-style-type: none"> Vue d'ensemble différenciée des 250 dernières années et de l'histoire antérieure (chronologie, tenants et aboutissants): 8 enseignants en histoire, 5 enseignants d'autres filières d'études d'orientations diverses; amples connaissances de l'histoire de la Suisse: 1 enseignant en histoire. Conscience de l'historicité: 1 enseignant en germanistique et 1 enseignant en histoire. Perception exempte de poncifs de l'histoire du Moyen-Age: 2 enseignants en germanistique. Connaissances de l'Antiquité (3). Connaissances de l'histoire américaine et de l'époque coloniale: 1 enseignant en anglais. Education civique (3).
Economie & droit	<ul style="list-style-type: none"> Compréhension générale de l'économie et de ses structures fondamentales: 3 enseignants en histoire. Fondements du droit: 1 enseignant en pédagogie. Connaissances concernant les coûts, intérêts et titres: 1 enseignant en informatique.
Géographie	<ul style="list-style-type: none"> Connaissances générales: 1 enseignant en histoire.
Chimie	<ul style="list-style-type: none"> Connaissances actuelles, formation préparatoire en sciences: 1 enseignant en médecine.
Physique	<ul style="list-style-type: none"> Acoustique: 1 enseignant en germanistique et 1 enseignant en anglais. Maniement des unités de mesures, connaissances concernant la pression, la pesanteur et les vitesses momentanées: 1 enseignant en informatique et 1 enseignant en architecture.
Philosophie	<ul style="list-style-type: none"> Connaissances fondamentales, connaissance des récits bibliques: 5 enseignants en histoire, 3 autres enseignants de filières d'études d'orientations diverses. Philosophie antique: 1 enseignant en pédagogie. Eléments de base des techniques argumentatives: 1 enseignant en sciences politiques.
Informatique	<ul style="list-style-type: none"> Connaissances d'utilisateur (6). Gestion des algorithmes: 1 enseignant en anglais. Courage de programmer, c'est-à-dire connaissance de langues de programmation: 1 enseignant en sociologie et 1 enseignant en informatique.
Musique	<ul style="list-style-type: none"> Aucune lacune n'est relevée.
Arts visuels	<ul style="list-style-type: none"> Connaissance de l'analyse des représentations visuelles: 1 enseignant en anglais.

Dans presque toutes les disciplines, des lacunes dans les connaissances et capacités initiales des étudiants sont mentionnées. Les lacunes relevées sont particulièrement nombreuses dans la langue première et les mathématiques. S'agissant de la langue première, les lacunes concernent pour l'essentiel également les composantes de la maîtrise de la langue, dont l'«indigence» est surtout critiquée par les germanistes. En l'occurrence également, seuls quelques germanistes et un angliciste ont relevé des lacunes dans les contenus essentiels de la germanistique.

3.3.4 Importance des compétences interdisciplinaires pour l'enseignement (question 4)

Les enseignants ont dû, pour répondre à la question suivante, évaluer sur une échelle de sept degrés l'importance qu'ils accordent aux compétences interdisciplinaires présentées quant à la compréhension et à la fréquentation fructueuse de leurs enseignements. La question posée était formulée comme suit:

«De quelles compétences interdisciplinaires supplémentaires importantes doivent disposer les étudiants dès le début de leurs études pour être en mesure de suivre votre enseignement?»

Les compétences interdisciplinaires qui apparaissent à la première colonne du tableau II.9 ont été pour l'essentiel reprises de l'enquête de Notter & Arnold (2003 et 2006) (cf. chapitre I, section 3), mais certaines ont été modifiées et complétées en fonction d'autres composantes de l'aptitude générale aux études, citées dans la littérature, et sur la base de discussions internes au sein de l'équipe EVAMAR (cf. chapitre I, section 5.4).

Huber (1998, p. 158) décrit par exemple les compétences générales et les attitudes suivantes:

- (par rapport à l'objet) motivation de l'étudiant, volonté d'apprendre, intérêt pour la discipline, curiosité;
- (par rapport à l'avenir) orientation, certitude dans le choix de la discipline et dans l'optique professionnelle (ou abstraction provisoire consciente de celle-ci);
- (par rapport à la sociabilité) volonté et capacité de contact, de communication et de coopération;
- (par rapport à soi-même) conscience réaliste de soi-même ou davantage encore: confiance en soi et force suffisante du moi pour se mettre en question, identité développée ou en développement et, en résultant:
- autonomie: quant aux contenus (détermination des thèmes), organisationnelle (espace, temps, coordination), méthodologique (techniques de travail scientifiques), logistique et pratique (habitation, travail).»

Schüpbach et al. (2006, 156) ont identifié les éléments suivants comme composantes d'une compétence opérationnelle dans le cadre des études d'anglais au niveau universitaire: acquisition autonome de la pratique de la langue anglaise, recherche d'information, travail d'équipe, apprentissage autogéré, utilisation des ressources, préparation des examens, limitation du temps libre, travail autonome, participation active aux enseignements. On trouve des résultats semblables chez Wetzenstein (2004) à partir de l'analyse des exigences posées à des études en psychologie: les qualités personnelles et les aptitudes prioritaires sont l'auto-organisation, la compétence linguistique générale, la motivation à performer, la pensée logique, la pensée critique, l'intelligence, la résistance aux sollicitations, la capacité à mémoriser, l'intérêt pour l'inconnu/l'ouverture et la pensée abstraite.

Les compétences partielles utilisées dans le questionnaire représentent pour la plupart des concepts encore assez peu fondés théoriquement et empiriquement. Mais l'intégration de l'enquête auprès des enseignants permet la comparaison avec l'étude, mentionnée ci-dessus, dans laquelle les données empiriques relatives à l'importance des compétences partielles ne sont relevées que sur la base de l'estimation des étudiants et non pas du point de vue des enseignants d'université.

Les résultats sont présentés au tableau II.9, classés selon les valeurs moyennes (la compétence est jugée d'autant plus importante que la moyenne est élevée).

Tableau II.9: Importance des compétences interdisciplinaires pour le propre enseignement

Compétences	Moyenne	Ecart-type	Importance du point de vue des étudiants*
Apprentissage autonome	6.48	.828	1
Prise de responsabilité pour son propre apprentissage et son propre travail	6.40	.872	3
Travail autonome	6.38	.882	2
Pensée critique	6.33	1.003	7
Capacité de résoudre les problèmes	6.11	1.243	6
Capacité d'expression écrite	5.89	1.407	8
Allocation efficace du temps et fixation des priorités	5.85	1.104	4
Gestion des sollicitations	5.59	1.339	5
Compréhension auditive dans le cadre des cours	5.55	1.420	
Capacité d'expression orale	5.42	1.715	10
Préparation d'amples examens	5.25	1.586	
Développement de nouvelles idées	5.20	1.374	11
Traitement d'une matière d'apprentissage volumineuse en peu de temps	5.14	1.405	
Capacité de recherche	5.02	1.648	
Travail en équipe	4.74	1.544	9
Présentation en public	4.16	1.819	12

N max. = 140

* Rang d'importance les premiers semestres du point de vue des étudiants selon Notter & Arnold (sans la catégorie «compréhension de texte») (2006, p. 52).

Toutes les compétences présentées ont donc été évaluées comme très importantes (proches de 7) à au moins moyennement importantes (légèrement supérieures à 4). L'ordre du classement tend à correspondre étonnamment bien avec l'évaluation de l'importance de ces compétences interdisciplinaires par les étudiants dans les études de Notter & Arnold (2003 et 2006). Globalement, la comparaison de cette évaluation par disciplines d'études confirme les degrés d'importance attribués. Seule la compétence «présentation en public» est inférieure à 4 dans de nombreuses filières d'études.

Les enseignants ont eu la possibilité d'indiquer, dans une partie de réponse ouverte, d'autres compétences interdisciplinaires qu'ils jugent importantes. Si certaines des aptitudes citées sont proches ou découlent directement de celles présentées, d'autres peuvent être considérées comme des compléments judicieux: la capacité de concentration, l'endurance, la précision, la pensée logique, la pensée en réseau, la pensée abstraite (par ex. la capacité de concevoir en trois dimensions), la pensée analytique, la pensée associative, la capacité de transfert, l'appréciation réaliste de soi, l'autocritique, la capacité argumentative, l'aptitude à discuter, la compétence de lecture.

3.3.5 Carences constatées dans les connaissances et capacités interdisciplinaires (question 5)

On a également demandé, pour les compétences interdisciplinaires, quelles carences ont été relevées. La question posée était formulée comme suit:

«Parmi les connaissances et capacités interdisciplinaires jugées très importantes à importantes sous la question 4 ci-dessus (valeurs 1-3), quelles sont celles, qui permettraient de suivre votre enseignement, dont les étudiants ne disposent généralement pas (plus) au début de leurs études?»

Les réponses à cette question ouverte ont été attribuées aux catégories présentées à la question 4 et quantifiées (cf. tableau II.10). On est frappé par la forte proportion des carences mentionnées concernant la capacité d'expression écrite, la pensée critique, le travail autonome et l'apprentissage autonome. Or, ces compétences interdisciplinaires occupent justement les degrés supérieurs de l'échelle d'importance. Cette corrélation entre l'importance de la compétence et l'indication d'une carence (plus la compétence est importante, plus elle fait défaut) s'observe sur l'ensemble du tableau. Par ailleurs, elle correspond fort bien avec l'autoévaluation des étudiants dans l'étude Notter & Arnold (2006).

Tableau II.10: Carences dans les connaissances et capacités interdisciplinaires jugées importantes

Compétences	Pas d'insuffisance	Compétence importante, mais lacunaire	Moyenne de l'importance
Capacité d'expression écrite	57.6%	42.4%	5.89
Pensée critique	65.0%	35.0%	6.33
Travail autonome	70.0%	30.0%	6.38
Apprentissage autonome	74.0%	26.0%	6.48
Capacité d'expression orale	79.0%	21.0%	5.42
Prise de responsabilité pour son propre apprentissage et son propre travail	79.4%	20.6%	6.40
Capacité de résoudre les problèmes	80.0%	20.0%	6.11
Capacité de recherche	81.0%	19.0%	5.02
Allocation efficace du temps et fixation des priorités	84.7%	15.3%	5.85
Gestion des sollicitations	84.8%	15.2%	5.59
Compréhension auditive dans le cadre des cours	88.9%	11.1%	5.55
Développement de nouvelles idées	91.0%	9.0%	5.20
Préparation d'amples examens	91.8%	8.2%	5.25
Traitement d'une matière d'apprentissage volumineuse en peu de temps	92.0%	8.0%	5.14
Présentation en public	97.0%	3.0%	4.16
Travail en équipe	97.0%	3.0%	4.74

N max. = 100

3.3.6 Connaissances et capacités de départ supplémentaires souhaitables

Par la dernière question du questionnaire que voici, on a demandé aux enseignants quelles seraient les connaissances et capacités supplémentaires qu'ils souhaiteraient voir dans les curricula du gymnase, au-delà des éléments indispensables pour suivre leur propre enseignement.

«Si vous pouviez fixer vous-même les objectifs de la formation gymnasiale, quelles seraient selon vous – comparativement à vos réponses aux questions 2 et 4 – les connaissances et capacités spécifiques à une discipline et

interdisciplinaires supplémentaires que les étudiants devraient avoir déjà acquis au gymnase en vue de votre enseignement? (merci de justifier votre réponse)»

Les résultats sont récapitulés au tableau II.11⁵. Ils ne sont que partiellement cohérents en ce sens qu'ils contiennent des éléments qui sont enseignés aujourd'hui déjà dans la plupart des gymnases (ce point apparaît particulièrement pour les mathématiques). On peut donc supposer que nombre d'enseignants sont insuffisamment informés des matières effectivement enseignées au gymnase.

Tableau II.11: Connaissances et capacités supplémentaires souhaitables au début des études

Branche	Contenus
Langue première	<ul style="list-style-type: none"> • Concepts utiles à l'analyse et à l'interprétation de textes et d'énoncés littéraires: 4 enseignants en germanistique, 1 enseignant en anglais. • Interprétation textuelle sous l'angle de son propre vécu et de ses propres valeurs, observation du texte: 1 enseignant en germanistique. • Connaissances de l'histoire littéraire: 1 enseignant en germanistique et 1 enseignant en pédagogie. • Connaissance de la littérature allemande et de ses parallèles avec l'histoire de la littérature anglaise: 1 enseignant en anglais. • Rhétorique: 1 enseignant en germanistique. • Description du contexte historique des textes, conscience des niveaux de langue et de l'évolution de la langue au fil de l'histoire: 1 enseignant en germanistique. • La langue première comme instrument de communication professionnelle: 1 enseignant en droit. • Connaissances de la linguistique, langue des jeunes et des médias, transformation de la langue, systèmes de signes, structures linguistiques: 1 enseignant en germanistique et 1 enseignant en anglais. • Maîtrise de la langue (3); aptitude à s'exprimer: 1 enseignant en sciences politiques et 1 enseignant en droit. • Rédaction de divers types de textes: 1 enseignant en germanistique; écriture structurée, scientifique, axées sur la discipline visée: 1 enseignant en informatique. • Grammaire, syntaxe, connaissance du texte, orthographe, ponctuation, approche sans appréhension des textes anciens: 1 enseignant en germanistique.
Mathématiques	<ul style="list-style-type: none"> • Calcul des probabilités: 1 enseignant en psychologie et 1 enseignant en physique; stochastique: 1 enseignant en pédagogie. • Mise en œuvre des théorèmes dans des problématiques pratiques concrètes: 1 enseignant en sociologie. • Introduction de la notion d'équation différentielle: 1 enseignant en informatique. • Démonstration par induction: 1 enseignant en informatique; principes de démonstration: 1 enseignant en physique. • Compréhension intuitive des concepts géométriques et analytiques: 1 enseignant en physique. • Pensée algorithmique: 1 enseignant en physique. • Compréhension des statistiques, présentation parlante des données, interprétation et analyse: 1 enseignant en biologie. • Nombres complexes: 1 enseignant en biologie. • Equations différentielles simples: 1 enseignant en biologie. • Produit vectoriel, équation circulaire, fonctions logarithmiques: 1 enseignant en architecture. • Compréhension de relations fonctionnelles à partir de représentations graphiques: 1 enseignant en architecture.
Biologie	<ul style="list-style-type: none"> • Bases biologiques des comportements, neurophysiologie et physiologie des organes sensoriels: 1 enseignant en psychologie. • Traitement d'un plus grand nombre de problématiques écologiques, connaissances des problèmes écologiques actuels en Suisse et dans le monde (3). • Aspects historiques de la biologie, biologie comparée: 1 enseignant en biologie. • Connaissances des organismes en relation avec la théorie de la biodiversité: 1 enseignant en biologie et 1 enseignant en médecine. • Cycles des substances, évolution: 1 enseignant en médecine.

⁵ Si aucune indication supplémentaire n'accompagne le chiffre, les mentions proviennent d'enseignants de diverses disciplines.

Branche	Contenus
Grec	<ul style="list-style-type: none"> • Connaissance de l'alphabet grec: 1 enseignant en histoire. • Connaissances fondées en langues anciennes: 1 enseignant en histoire.
Latin	<ul style="list-style-type: none"> • Connaissances minimales du vocabulaire de base: 1 enseignant en germanistique. • Les connaissances de latin devraient être obligatoires: 3 enseignants en histoire. • Formation de la sensibilité linguistique en langue première, meilleure maîtrise des concepts médicaux: 1 enseignant en médecine.
Français	<ul style="list-style-type: none"> • Compréhension sans difficulté des textes et débats en français: 1 enseignant en sciences politiques.
Anglais	<ul style="list-style-type: none"> • Anglais au niveau de l'«advanced level» (au moins C2): 2 enseignants en anglais. • Traitement également de types de texte difficiles (langage scientifique): 1 enseignant en anglais. • Connaissances de base de l'histoire à intégrer dans l'histoire de la littérature anglaise: 1 enseignant en anglais. • Compétence linguistique plus élevée: 1 enseignant en anglais. • De bonnes connaissances (3).
Italien/espagnol	<ul style="list-style-type: none"> • Lire la littérature spécialisée en italien: 1 enseignant en histoire.
Histoire	<ul style="list-style-type: none"> • Examen des processus de l'histoire: 1 enseignant en germanistique; enseignement de philosophie conjoint: 1 enseignant en germanistique; connaissances en histoire: 1 enseignant en anglais. • Rudiments de connaissance de la structure des événements survenus en Suisse et en Europe: 1 enseignant en histoire; connaissances des faits établis et panoplie basique de l'historien: 1 enseignant en histoire; connaissance des importantes époques de l'histoire mondiale et de ses dates: 1 enseignant en histoire; connaissances générales sur les bases sociales, culturelles et politiques de son propre «milieu de vie»: 1 enseignant en histoire; large panorama des événements et problématiques de l'histoire européenne: 2 enseignants en histoire; l'histoire a commencé avant 1789: 2 enseignants en histoire; connaissance des relations de l'histoire mondiale après la Révolution françaises: 1 enseignant en sciences politiques; aperçu de l'histoire des 300 dernière années: 1 enseignant en droit. • Trois volets du concept de politique: la forme («polity»), les procédures («politics»), les contenus («policy»): 1 enseignant en sciences politiques. • Sens de l'orientation conférée par l'étude de l'histoire comme base d'appréciation de l'époque actuelle: 1 enseignant en sciences politiques. • Bases élémentaires d'éducation civique (structure de la Confédération, Constitution, pouvoirs de l'Etat): 1 enseignant en droit. • Histoire de la culture et des techniques: 2 enseignants en architecture. • Histoire des médias: 1 enseignant en sciences de la communication.
Economie & droit	<ul style="list-style-type: none"> • Connaissances économiques des médias: 2 enseignant en sciences de la communication; en outre, connaissances juridiques concernant les médias: 1 enseignant en sciences de la communication. • Principes et modèles microéconomiques: 1 enseignant en économie. • Bases de la gestion des RH: 1 enseignant en économie. • Droit: principaux aspects juridiques utiles dans la vie quotidienne: 1 enseignant en économie.
Géographie	<ul style="list-style-type: none"> • Ethnologie: 1 enseignant en histoire.
Chimie	<ul style="list-style-type: none"> • Etablir la chimie comme discipline scolaire obligatoire: 1 enseignant en médecine; réduction des préjugés envers la discipline (environnement): 1 enseignant en médecine.
Physique	<ul style="list-style-type: none"> • Pas de communication.
Pédagogie, psychologie, philosophie	<ul style="list-style-type: none"> • Enseignement conjoint de la philosophie et de l'histoire: 1 enseignant en germanistique. • Connaissances de base permettant d'intégrer les phénomènes relevant de la recherche littéraire dans les études d'anglais: 1 enseignant en anglais. • Etude comparée des religions ou introduction à la pensée et au travail théologique: 1 enseignant en histoire; connaissances de l'histoire et de la culture compte tenu de la relation entre la société et la religion: 1 enseignant en histoire. • Philosophie: théorie de la connaissance comme introduction aux méthodes de travail des sciences sociales: 1 enseignant en sciences politiques. • Etendre la psychologie à la sociologie, compréhension du fait que la réalité est construite par l'être humain: 1 enseignant en architecture.

Branche	Contenus
Informatique	<ul style="list-style-type: none"> • Connaissances utilisateur approfondies en informatique: 1 enseignant en germanistique; connaissances utilisateur en informatique: 1 enseignant en pédagogie; utilisation courante du PC et des systèmes: 1 enseignant en sociologie. • Recours aux nouveaux médias: 1 enseignant en germanistique. • Connaissances fondamentales de la programmation: 1 enseignant en biologie.
Arts visuels	<ul style="list-style-type: none"> • Connaissances culturelles: 1 enseignant en anglais.
Musique	<ul style="list-style-type: none"> • Large culture générale: 1 enseignant en médecine.
Aptitudes interdisciplinaires	<ul style="list-style-type: none"> • Techniques et stratégies d'apprentissage (3). • Mise en réseau des contenus appris (3). • Préparation à l'apprentissage permanent: 1 enseignant en germanistique. • Technique de citation scientifique (3). • Collecte d'information et recherche (5); traitement critique des informations (4). • Rédaction de travaux analytiques, développement complet et présentation de ses propres thèses: 2 enseignants en anglais et 2 enseignants en germanistique. • Ponctualité: 1 enseignant en germanistique. • Précision du travail (3). • Motivation personnelle accrue: (4); tolérance à la frustration: 1 enseignant en informatique et 1 enseignant en biologie. • Capacité de conduire une réflexion critique (6); pensée interdisciplinaire: 2 enseignants en anglais; distinction des faits et des avis: 2 enseignants en pédagogie; pensée structurée autonome et pensée logique: 1 enseignant en économie. • Davantage de pratique d'écriture (3). • Lecture approfondie et capacité de saisir les structures textuelles abstraites, complexes et profondes: 4 enseignants en histoire et 4 autres enseignants de diverses filières d'études. • Interprétation des représentations visuelles: 1 enseignant en histoire. • Amélioration des techniques de présentation: 3 enseignants en histoire. • Travail autonome (3). • Apprentissage autonome (3). • Etude des principes de la connaissance scientifique: 1 enseignant en pédagogie; intérêt pour le travail scientifique: 1 enseignant en économie; de manière générale, plaisir à découvrir: 1 enseignant en psychologie. • Compétences sociales, travail en équipe: 1 enseignant en psychologie et 1 enseignant en histoire. • Aptitudes à communiquer: 1 enseignant en biologie. • Compétences organisationnelles: 1 enseignant en architecture et 1 enseignant en biologie. • Capacité d'expression orale et écrite: 1 enseignant en architecture.

3.3.7 Conclusions

Les résultats de l'enquête n'ont presque sans exception pas appelé de correction des résultats obtenus des analyses A1 et A3 en vue des tests, mais ils sont intéressants en raison de leur caractère de résultats autonomes du rapport principal. Notons toutefois les exceptions suivantes: un domaine «compétences en grammaire et en orthographe» a été intégré au test pour la langue première. Par un questionnaire complémentaire, on a demandé aux gymnasiens d'autoévaluer les compétences interdisciplinaires importantes. Enfin, les résultats de l'enquête confortent le choix de la biologie comme discipline scientifique expérimentale (cf. chapitre IV).

4 Sous-projet A2: analyse des examens d'entrée

4.1 Aperçu

Le présent sous-projet devait permettre, à l'aide d'une analyse des examens d'entrée dans les universités suisses, de déterminer des indications complémentaires quant aux compétences requises pour des études universitaires.

Ces examens comprennent avant tout les examens d'admission proposés dans la plupart des universités comme possibilité de se qualifier pour la filière d'étude, lorsque la preuve de l'existence d'une maturité suisse fait défaut. Dans le cadre de la collecte des données pour les sous-projets A1 et A3, on a également demandé aux universités contactées de fournir des indications sur d'éventuels examens d'admission. En complément, nous avons aussi consulté à ce sujet les pages internet de toutes les universités suisses.

Un premier aperçu a révélé que ces examens sont généralement couplés au programme des examens fédéraux de maturité et qu'ils sont donc très semblables aux examens passés dans le cadre du gymnase. Ces derniers étant l'objet de l'analyse réalisée dans le sous-projet D1, cette catégorie d'examen d'admission a donc été écartée de la suite de l'étude pour parer au risque d'établir des clés d'interprétation circulaires.

La seconde catégorie d'examen d'entrée est formée des tests d'aptitude à proprement parler. Ces tests sont étayés par des méthodes de prévision de la réussite dans les études explicitement formulées, dotées d'une assise théorique et, dans la mesure du possible, validées empiriquement. Les renseignements pris auprès des diverses universités ont montré que seul le test d'aptitudes pour les études de médecine (AMS) et l'examen d'admission de l'Université de Saint-Gall pour les candidats étrangers aux études entrent dans cette catégorie. Nous avons ensuite examiné ces deux instruments de plus près.

4.2 Le test d'aptitude aux études de médecine

Le test d'aptitudes pour les études de médecine (AMS) est utilisé en Suisse par diverses universités dans le cadre de la sélection de l'accès aux études de médecine en raison du nombre limité des places d'études. Il contient des tâches qui permettent de saisir les compétences interdisciplinaires. Il s'agit d'un test d'aptitude bien établi sur le plan théorique, dont la validité des prévisions est optimale dans le cadre des possibilités. Il se fonde sur le TMS, le test développé en Allemagne pour les filières d'études médicales («Test für medizinische Studiengänge»).

L'AMS est un test d'aptitudes axé sur le domaine des sciences expérimentales. Hänsgen & Spicher (2007, 12) le décrivent de manière synthétique comme suit: «On n'y demande pas des faits, mais on y donne des problématiques à résoudre en fournissant les informations nécessaires. L'objectif est de soumettre à une épreuve de travail pour mesurer l'aptitude aux études. Les analyses des exigences réalisées en son temps pour les études de médecine, qui ont requis un lourd investissement, ont été exécutées sous divers angles. Sur la base de ces analyses, quelque 50 domaines et types de tâches ont été examinés. 13 d'entre eux, qui sont apparus les plus appropriés sous de nombreux aspects, ont été retenus pour les essais de la procédure transitoire du TMS et testés empiriquement. Compte tenu surtout des résultats de validité obtenus, 9 groupes de tâches ont été intégrés dans le TMS. L'AMS offre une très bonne prévision de l'aptitude aux études en la mesurant selon des critères comparables à ces tests. La corrélation des résultats des tests avec les notes des examens durant les études se situe dans un domaine moyen (cf. les chiffres concrets dans Hänsgen & Spicher, 2001). Les sous-tests pour lesquels ces chiffres sont disponibles, sont les suivants.

- Problèmes quantitatifs et formels: capacité de «gérer correctement des chiffres, des grandeurs, des unités et des formules dans le cadre de problématiques médicales et scientifiques» (Hänsgen & Spicher, 2007, p. 97).
- Travail concentré et soigneux: «capacité de travailler rapidement, soigneusement et de manière concentrée» (p. 101).
- Compréhension de questions fondamentales de la médecine et des sciences naturelles: «compréhension des questions de médecine et des sciences expérimentales...». Le texte pourrait être tiré d'un manuel. Pour

comprendre ce passage de texte, il importe de savoir si des conclusions logiques déterminées peuvent en être déduites. Tous les faits nécessaires pour s'acquitter de la tâche se trouvent dans le texte. En particulier, des connaissances préalables de médecine ne sont pas requises» (p. 101).

- Reconnaissance de fragments de figures: capacité «de reconnaître des fragments dans une représentation visuelle complexe» (p. 103).
- Diagrammes et tableaux: capacité «d'analyser et d'interpréter correctement des diagrammes et des tableaux. Pendant les études, de nombreuses relations d'interdépendances seront représentées sous cette forme» (p. 104).
- Compréhension de textes: capacité «d'absorber et de traiter un matériel textuel volumineux et complexe» (p. 98).
- Mémorisation de faits: capacité d'enregistrer les faits concernant les patients et de s'en souvenir (p. 103).
- Figures tubulaires, capacité de représentation spatiale: «une fonction qui est importante par exemple pour comprendre les radiographies. Pendant les études, nombre de structures et processus tridimensionnels seront en fait présentés sous la forme de représentations en deux dimensions» (p. 97).
- Mémorisation de figures: examine dans quelle mesure on peut s'imprégner et noter les détails d'objets. (p. 102).

On trouvera des tâches types par exemple dans Hänsgen & Spicher (2007), où se trouve décrit le sous-test «Planifier et organiser», qui est actuellement utilisé.

4.3 L'examen d'admission à la HSG

L'institut qui a développé la procédure d'examen d'admission appliquée par l'Université de Saint-Gall pour les candidats étrangers aux études est celui-là même qui avait élaboré le test pour les filières médicales en Allemagne (TMS). Contrairement à l'AMS, on ne dispose toutefois pas encore d'étude empirique sur sa validité prévisionnelle. Une ossature théorique fait également défaut.

4.4 Conclusions

Les tests prévus pour le module C selon le mandat ont le caractère de tests des connaissances scolaires standardisés⁶, qui revêtent déjà, selon l'estimation de Schallberger (1996), une valeur prévisionnelle élevée. Comme l'insertion de tâches issues d'un véritable test d'aptitude aux études offrait la possibilité de mesurer, en complément des tests spécifiques prévus dans EVAMAR II, des compétences pertinentes pour les études, mais plutôt indépendantes des compétences cognitives, le groupe de pilotage du projet EVAMAR II a décidé de présenter également, dans le cadre du module C réalisé en Suisse alémanique, des tâches provenant de versions antérieures du test d'aptitudes pour les études de médecine (AMS). Cette partie du test n'était pas prévue dans le projet de mandat.

⁶ Pour la langue première, d'autres considérations ont conduit au développement d'une grille de compétences et d'un test pour des capacités linguistiques interdisciplinaires plus générales (cf. chapitre III, section 2).

III Module B: Modélisation des connaissances et capacités requises pour entreprendre des études universitaires dans certaines disciplines

1 Généralités

La planification d'EVAMAR II prévoyait initialement de condenser directement en modèles de compétence les résultats des analyses des sous-projets A1 à A4. La réalisation de ce projet n'a toutefois pas tardé à se révéler trop difficile pour diverses raisons, la principale étant que les textes de cours et exercices analysés dans le module A se prêtaient certes à établir des relations directes avec les connaissances et compétences absolument indispensables à une appréhension aisée, mais ne fournissaient pas assez d'informations susceptibles de composer un modèle de compétence cohérent, correspondant à l'état actuel des connaissances théoriques en matière d'apprentissage et validé empiriquement, selon la conception prônée actuellement dans les discussions théoriques autour des standards de formation (cf. surtout Weinert, 2001). De surcroît, les unités de sens ou de connaissance et les capacités cognitives ainsi établies ne recouvraient pas totalement les exigences des universités (voir détails au chapitre II). Il s'agissait seulement d'un échantillon, lui-même encore trop étroit pour servir de base unique à un modèle de compétence embrassant l'ensemble de l'aptitude aux études supérieures.

Il s'est donc avéré nécessaire d'analyser à la lumière des objectifs d'EVAMAR II la pertinence des modèles de compétence qui existaient dans chaque discipline, de les reprendre – intégralement ou en partie – ou de les perfectionner. Les pages qui suivent présentent brièvement les modèles analysés et indiquent leur degré de pertinence pour la problématique du projet EVAMAR II. Le module B ne visait pas à procéder à une modélisation des compétences sur la base d'une lecture «déterministe» des résultats du projet A; l'enjeu était bien plutôt de combiner cette grille de compétences avec des modèles déjà existants. Ces adaptations dûment documentées ont essentiellement servi à préciser, à l'aide des connaissances et compétences spécifiques identifiées dans le projet A, les descriptions plutôt générales et formelles des domaines couverts par les modèles de compétence, dans l'optique de l'interface qu'étudie EVAMAR II. Des modèles ont donc été développés pour les trois disciplines testées dans le module C et étudiées dans D1: la langue première, les mathématiques et la biologie. Pour respecter la définition de la notion, ou plutôt du concept construit de «modèle de compétence» attesté dans les discussions actuelles et qui a une vocation générale au sens de Weinert (2001) ou de Klieme et al. (2006), nous avons appelé «grilles de compétence» les résultats de notre modélisation. Celles-ci se présentent comme une matrice avec les dimensions des «contenus» et des «compétences»: ces dernières s'y trouvent notées comme des descripteurs de capacité qui font état des savoirs et des savoir-faire dans chacun des domaines de la discipline analysée.

A la différence des modélisations de compétences actuellement déjà existantes ou en préparation, le projet EVAMAR II requérait, pour les disciplines considérées, une attention toute particulière aux cinq aspects suivants de l'orientation du projet:

1. le projet est centré sur des degrés relativement avancés - il considère de jeunes adultes de 18 à 19 ans peu avant les examens finaux du gymnase, par conséquent sur des niveaux de compétences plutôt élevés;
2. l'accent est mis sur le type d'école «gymnase»;
3. l'étude est axée centrée sur l'appréciation de l'aptitude aux études supérieures, ou maturité universitaire en général (objectif de la formation gymnasiale) peu de temps avant un passage possible à l'une des hautes écoles du pays;
4. l'analyse est centrée sur la modélisation et la vérification de savoirs et savoir-faire prérequis pour les premiers semestres des études selon les données recueillies dans le projet A;

5. l'analyse est centrée sur la modélisation et la mesure de performances sous forme d'items élaborés dans le cadre du projet C d'EVAMAR II.

Nous avons obtenu de cette manière, pour les trois disciplines soumises aux tests, des champs de compétence s'accordant avec les connaissances et compétences identifiées dans le projet A. Le module C (cf. chapitre IV du rapport) a ensuite construit pour chacun de ces champs des items se rapportant aux contenus spécifiques définis dans le projet A. Tous les items du test peuvent donc se référer aussi bien à la modélisation choisie qu'aux supports de cours universitaires. Ce «tramage des points d'intersection» concret et systématique opéré par le test ou, plus exactement, par la mesure des performances effectuée dans le projet EVAMAR II, n'avait encore jamais été réalisé dans un contexte semblable, c'est donc une première.

2 La grille de compétences en langue première

2.1 Bases

En comparaison avec les autres domaines de compétences que le projet EVAMAR II analyse pour les modéliser, celui de la langue première présente les spécificités suivantes:

- La base de données générée dans le sous-projet A1 pour la langue première – y compris le fonds de vocabulaire emprunté aux langues d’origine comme le latin – est alimentée très largement par toutes les disciplines d’études analysées (cf. chapitre II, section 2.4). En conséquence, la grille des compétences fondée sur cette base ne répond pas aux exigences du projet EVAMAR II, car il s’agit d’un modèle fortement lié à un cursus de discipline, avec une orientation claire vers les études de linguistique et qui met l’accent sur des contenus spécifiques de l’enseignement de la langue première comme les connaissances préalables d’histoire de la littérature, de sciences littéraires, de technique de l’analyse de textes, etc. Or, l’analyse d’EVAMAR II doit bien davantage avoir en point de mire les compétences langagières générales acquises en langue première (réception, production, conscience du fonctionnement de la langue).
- Les discussions engagées dans la littérature spécialisée (par exemple Ehlich & Steets, 2003; Feilke & Steinhoff, 2003 et Kruse, 2007 sur la compétence rédactionnelle *Schreibkompetenz*; vue d’ensemble et bibliographie annotée chez Mugabushaka, 2005) ainsi que nos recherches menées dans les sous-projets du module A s’accordent pour constater qu’un niveau suffisant de compétences de communication en langue première – au sens des compétences langagières, sociolinguistiques et pragmatiques⁷ – est indispensable à une maîtrise des études universitaires quelles qu’elles soient (cf. chapitre II) et que de «solides qualifications clés en germanistique» *«germanistisch fundierte Schlüsselqualifikationen»* (Mugabushaka, 2005, p. 62) doivent sans cesse être optimisées dans les hautes écoles. Ces qualifications englobent un vocabulaire large et centré sur l’emploi scientifique, en particulier sur le genre des textes «spécialisés universitaires» (*«universitäre Fachtexte»*, par exemple Ehlich, 2003; Gläser, 1990, cité selon Thim-Mabrey, 2005). En comparaison, la capacité de faire appel à des connaissances spécifiques de littérature joue un rôle secondaire. Ce savoir littéraire de base dans l’enseignement de la langue première est utile dans divers contextes; pourtant, comme l’a montré le module A, il n’est pas un prérequis indispensable, même pour les études de germanistique notamment.
- La grille de compétences à développer devrait pouvoir s’appliquer aux trois langues premières en usage en Suisse à la maturité: l’allemand, le français et l’italien.

On citera d’abord ici le projet de recherche qui a donné lieu à la grille d’analyse zurichoise (*Zürcher Textanalyseraster*, Sieber, 1994), le projet de recherche mené sur l’allemand et l’anglais DESI (*Deutsch-Englisch-Schülerleistungen-International* DESI - DESI-Konsortium, 2006; Klieme, 2006; Beck & Klieme, 2007; Willenberg, 2007; DESI-Konsortium, 2008), le Programme international pour le suivi des acquis des élèves (PISA – *Programme international pour le suivi des acquis des élèves*, - Haider & Reiter, 2001, 2004; Prenzel, 2004; Schiefele et al., 2004), le Cadre européen commun de référence CECR (Conseil de l’Europe, 2001; Schneider & North, 1999); en Suisse, les travaux concernant la réforme de l’école en général, dont l’accord intercantonal sur l’harmonisation de la scolarité obligatoire HarmoS (HarmoS, CDIP, 2004a, 2004b, 2005, 2007, 2008; Maradan &

⁷ Les compétences langagières de communication comportent plusieurs catégories:

1. les compétences linguistiques proprement dites, qui englobent notre connaissance de la langue, (lexique, sémantique, orthographe, prononciation) et de ses règles de fonctionnement (grammaire, ponctuation, etc.),
2. les compétences sociolinguistiques, qui nous permettent de respecter les conventions sociales et les normes culturelles présidant à l’utilisation de la langue (dont, par ex, formes de politesse, tournures obligées, etc.),
3. les compétences pragmatiques, qui règlent l’emploi fonctionnel de l’utilisation de la langue et par lesquelles nous savons utiliser la langue de façon pertinente en respectant par exemple le changement de locuteur, le développement d’un thème, la cohérence et la connaissance des genres de textes (Conseil de l’Europe, 2001, p. 25; Klieme, 2003, p. 148).

Mangold, 2005; Sieber, 2006a, 2006b, 2006c) et enfin (la liste n'est cependant pas exhaustive) – les relevés centralisés de l'état des prestations en neuvième année d'école menés en 2004 dans le land de Rhénanie-du-Nord-Westphalie (Nordrhein-Westfalen - NRW *Zentralen Lernstandserhebungen* (Leutner, Wirth & Fleischer, 2005; Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2005a, 2005b, 2007, 2008).

Ces études empiriques menées au cours de ces dernières années ne poursuivent pas toutes les mêmes objectifs et sont peu homogènes quant à leurs structures (pour ce qui est des années scolaires concernées par exemple, ou des types de textes); elles présentent donc peu de zones de recoupement dans les domaines de compétences ni d'unité de critères quant aux échelles de niveaux. En outre, ces études d'évaluation se fondent sur des conceptions assez différentes de la compétence de lecture (Streblo, 2004, p. 275). En particulier, l'originalité relative – pour ne pas dire la nouveauté – de l'approche d'EVAMAR II était de tester dans ses relevés de performances des savoirs et des savoir-faire qui se révèlent pertinents pour les études universitaires du point de vue le plus proche des hautes écoles elles-mêmes; cette spécificité n'a fait que rendre plus difficile un recours actif à la pluralité des travaux divers de la littérature spécialisée.

2.2 Grilles et modèles de compétences existants pour la langue première

2.2.1 Grille zurichoise d'analyse de textes (Zürcher Textanalyseraster)

2.2.1.1 Brève présentation

Le projet mené dans le cadre du Fonds national il y a plus de dix ans sur les compétences des bacheliers en langue maternelle («Muttersprachliche Fähigkeiten von Maturanden und Studienanfängern in der Deutschschweiz», Sieber, 1994) a été à notre connaissance le premier travail qui ait examiné en même temps dans tout le pays les performances en langue première des bacheliers et des personnes commençant les études universitaires. Il a consigné l'état des lieux ainsi que les souhaits formulés, en l'occurrence les attentes que formulaient les professeurs et le personnel enseignant des écoles moyennes et des universités. Le relevé a été fait par le biais d'enquêtes et d'entretiens semi-structurés. La base des données recueillies pour l'analyse des compétences langagières était constituée de 800 travaux (dissertations – travaux écrits) de bacheliers des années 1988 à 1990 de dix écoles moyennes alémaniques, en l'occurrence il s'agissait surtout de textes issus de l'enseignement (en allemand, physique, biologie) ainsi que quelques compositions de l'épreuve écrite de maturité. Il s'ajoutait à cela un corpus limité de textes d'étudiants débutants en germanistique à l'Université de Zurich (p. 191-200). Pour saisir et analyser ces travaux, Sieber et son équipe ont élaboré une grille différenciée, document que l'on connaît sous le nom de *Zürcher Textanalyseraster* (Grille zurichoise d'analyse de textes, Sieber, 1994, p. 149-186). La perspective théorique de cet instrument qui permet d'émettre des assertions sur la qualité et les défauts de textes est fournie par une théorie linguistique (Nussbaumer, 1991, 1993, cité par Sieber, 1994, p. 143-148), laquelle conçoit le texte comme une unité globale et cohérente qui demande à être reconstruite par le récepteur à l'aide du savoir tant linguistique que global et pragmatique. Dans le sens d'une fiche de travail heuristique («*Im Sinne einer heuristischen Folie*», p. 43) Sieber et son équipe se servent du modèle en couches du linguistique Glinz (1987, p. 206-236, cité par Sieber, 1994, p. 41-43) qui considère au départ sept domaines de compétences, avec un degré de pertinence décroissant (cf. figure III.1).

Figure III.1: Modèles des compétences langagières de Hans Glinz, dit modèle par couches (1987, cité par Sieber, 1994, p. 41)

I	Domaine central de la constitution de la personne <construction de l'identité>
II	Modèle pour la construction de textes et stratégies langagières
III	Domaine sémantique: les significations et les structures signifiantes
IV	Domaine lexical et morphosyntaxique: les mots et les structures grammaticales formelles
V	Formes globales de la conduite de la voix: les sons en amont des signifiants de mots
VI	Système phonologique: la morphologie des syllabes, <prononciation et articulation>
VII	Système graphique: orthographe et ponctuation.

Sieber et son équipe (1994) interprètent les résultats de leurs analyses sur les compétences langagières (p. 308 ss.) à la lumière de ce modèle de compétences de Glinz, quelque peu aménagé cependant: ils réunissent les secteurs IV et VII pour en faire le domaine «les mots et leurs structures formelles/écriture» ainsi que les secteurs V et VI dans le domaine «conduite de la voix et prononciation». C'est à la lumière de ces cinq domaines de l'oral et de l'écrit qu'ils effectuent l'évaluation des compétences orales et écrites des bacheliers et des étudiants et étudiantes des premiers semestres.

2.2.1.2 Appréciation de la grille zurichoise de textes pour une utilisation dans EVAMAR II

Les descripteurs de la grille zurichoise de textes (Sieber, 1994) sont compatibles avec l'intention du projet EVAMAR II dans la mesure où les données recueillies concernent aussi, pour une large part, le degré gymnasial orienté vers le passage à l'université.

Cependant, les résultats du projet de Sieber, pour le domaine partiel des compétences touchant à l'écrit, se fondent sur des textes qui, de par leur longueur, se prêtent à une analyse d'aspects importants de la rédaction comme la cohérence, la structure et la conduite de la lecture. Comme nous le précisons plus loin, les travaux d'EVAMAR II nous ont conduits à renoncer dans le cadre des tests à la production de longs textes réclamant beaucoup de temps, comme des lettres, des dissertations ou des analyses, pour leur préférer des formats ouverts et plus courts. En outre, le deuxième domaine partiel inclus dans l'appréciation des enseignants et experts zurichois, celui de la capacité d'expression orale, n'entrait pas dans les objectifs de l'étude d'EVAMAR II. En conclusion, on a renoncé à utiliser la grille d'analyse de textes zurichoise comme cadre théorique de la grille de compétences d'EVAMAR II réservée à la langue première.

2.2.2 DESI

2.2.2.1 Brève présentation

Commandité par la Conférence des ministres de l'éducation des länder allemands dans le sillage de PISA, pour les années 2001 à 2006, le programme DESI («*Deutsch-Englisch-Schülerleistungen-International*») est le projet de recherche didactique le plus important mené jusqu'ici à l'échelle du pays entier. Afin de tester de façon représentative l'enseignement des langues sous tous ses aspects, pas moins de 10 640 jeunes gens de toutes les divisions scolaires issus de 219 écoles de tous les länder ont été soumis à des tests au début et à la fin de la neuvième année scolaire dans le but de dégager des modèles de compétences scientifiquement fondés et empiriquement attestés applicables à tous les domaines de compétences partiels de l'enseignement de l'allemand. A ce matériel s'ajoutent des leçons d'anglais enregistrées sur vidéo aux fins d'une modélisation des compétences. Les tests étaient construits sur des analyses des consignes liées aux cursus d'études en vigueur dans les länder, par exemple selon les types de textes pour ce qui est des tâches d'écriture ou selon les espèces de mots pour ce qui est de la conscience linguistique.

Le test DESI proposé pour l'allemand langue maternelle repose sur six domaines de compétence comportant de trois à cinq niveaux⁸ (de A à E): plus précisément trois niveaux pour chacun des domaines de la compétence rédactionnelle, du vocabulaire, de l'orthographe et de l'argumentation, quatre niveaux pour la compréhension de l'écrit et cinq dans le domaine de la conscience de la langue). Les constats publiés de façon centralisée en 2006 (DESI-Konsortium, 2006; Klieme, 2006; Beck & Klieme, 2007; Willenberg, 2007) permettent de dégager des descripteurs de «capacité de faire» pour trois domaines de compétence. On trouve ceux-ci dans la figure III.2 ci-dessous.

Figure III.2: Représentation – par EVAMAR – de trois des six domaines de compétence de DESI pour la langue première «Allemand» (DESI-Konsortium, 2006) avec les niveaux de A à E (n'y figurent pas: l'argumentation, le lexique, l'orthographe)

Niveau de compétence	Conscience du fonctionnement de la langue (tableau 3, complété, pour E par les descripteurs du texte ⁹ , p. 9)	Compréhension de l'écrit (tableau 2, p. 5)	Compétence rédactionnelle (Description p. 7)	
			Compétence partielle de sémantique pragmatique	Compétence partielle liée au système linguistique
E	Utilisation active d'un savoir formulable: user activement d'un savoir explicite (déclaratif) sur des phénomènes complexes tels le discours indirect allemand ou les formes du subjonctif.	<i>(Pas de niveau E)</i>	<i>(Pas de niveau E)</i>	<i>(Pas de niveau E)</i>
D	Contrôles grammatical et stylistique développés: reconnaître et corriger des erreurs stylistiques complexes, dissiper des ambiguïtés.	Exploitation de modèles mentaux intégrés: savoir définir les personnages principaux et leurs relations, les coordonnées de lieu, de temps, et le motif central.	<i>(Pas de niveau D)</i>	<i>(Pas de niveau D)</i>
C	Contrôle grammatical développé: reconnaître et corriger avec assurance des phénomènes grammaticaux même complexes; savoir faire un classement simple des notions grammaticales les concernant.	Lecture référencée: user activement d'un savoir général ou contextuel; pouvoir mettre en relation deux passages différents du livre afin d'éclaircir les thèmes ou de définir des causalités.	Les lettres du niveau C sont rédigées dans un style exigeant / soigné, utilisent un lexique varié et font apparaître une structure logique.	(Aucune indication)

⁸ Dans DESI, la subdivision du continuum des compétences en paliers se fait par le biais de la notion de *niveau de compétence*. Les auteurs de DESI renoncent à traduire le concept anglais de *level (proficiency level)* par *Stufe* pour éviter une connotation piagétienne qui serait hors de propos (Hartig, 2007, p. 86).

⁹ Les descripteurs portant sur la conscience de la langue publiés ultérieurement varient souvent (Beck & Klieme, 2007, p. 154-157); comme nous n'en disposons pas encore dans la phase de modélisation, nous renonçons à les présenter (exemple pour le niveau E (p. 156): «utilisation active du savoir déclaratif par un commentaire réfléchi, savoir-faire aussi mobilisable dans des conditions de grande complexité.»).

Niveau de compétence	Conscience du fonctionnement de la langue (tableau 3, complété, pour E par les descripteurs du texte ⁹ , p. 9)	Compréhension de l'écrit (tableau 2, p. 5)	Compétence rédactionnelle (Description p. 7)	
B	Consciens grammaticale et stylistique simples: savoir créer un style et une cohérence avec des contextes et des contenus aisés à comprendre.	Lecture centrée: former des inférences, en particulier sur la base d'une lacune entre deux phrases; savoir se concentrer sur l'élément précis qui permet de résoudre une difficulté sémantique ou logique de compréhension; coordonner les nombres figurant dans des textes.	Les textes (c.-à-d. lettres) de ce niveau permettent une communication écrite exempte de difficulté de compréhension entre l'auteur de la lettre et le destinataire, Même rédigées avec un lexique relativement modeste, les lettres à ce niveau ont au moins un semblant de structure qui permet de repérer leurs éléments centraux.	Les textes rédigés respectent dans une large mesure les règles et conventions d'orthographe et emploient des constructions de phrase moyennement complexes
A	Conscience grammaticale simple: reconnaître les fautes de grammaire grossières et savoir (en partie) les corriger.	Lecture élémentaire: reconnaître les unités essentielles porteuses du sens d'une phrase ou d'un paragraphe.	Les textes de ce niveau suivent plutôt les usages linguistiques du registre familier, comportent des erreurs de logique et ne contiennent pas tous les éléments essentiels qu'on attend dans une lettre.	Les textes comportent des fautes d'orthographe, de ponctuation et de syntaxe qui entravent la communication ou – faute de système linguistique – rendent même impossible la transmission efficace de l'information.

2.2.2.2 Appréciation du modèle de compétence de DESI pour son emploi éventuel dans EVAMAR II

Le modèle de compétence employé dans DESI et reconnu valable a été conçu pour élaborer des tests en neuvième année scolaire et se réfère aux contenus des cursus d'enseignement des 16 länder allemands. Une transposition de ces conditions cadre pour la 12^e année s'imposait, c'est-à-dire pour les exigences universitaires et les catégories de textes spécialisés fréquents au niveau universitaire. Il eût été indispensable aussi de différencier le modèle DESI en employant surtout les niveaux supérieurs de l'échelle.

Au vu de nos objectifs, il était peu profitable de travailler sur le modèle DESI pour le domaine de compétence de la «production de textes» à partir du genre épistolaire utilisé par DESI (qui distingue aussi, il est vrai, les «lettres documentaires» et la correspondance personnelle à parts égales). En effet, une comparaison des types de textes en fonction des buts d'EVAMAR était peu pratique dès le départ ou alors faisable dans de rares cas, car il aurait fallu étendre les modèles de textes demandés à des compositions ou dissertations à sujet imposé ainsi qu'à des travaux de (pro)séminaire.

S'agissant du domaine de compétence intéressant de l'argumentation, il s'est avéré après examen approfondi comme désavantageux pour EVAMAR II que ce savoir faire ait déjà été pris en considération dans DESI, mais au titre de compétence d'analyse de formes et d'intentions argumentatives, c'est-à-dire dans la partie consacrée aux performances de reconnaissance, donc aux activités exclusivement réceptives.

Une autre complication nous a influencés: au moment de prendre les décisions pour le projet EVAMAR II, nous ne disposions pas encore des publications complémentaires de DESI – parues en 2007 et 2008 avec les modèles implémentés dans l'intervalle pour les performances des élèves et les explications complètes (déductions théoriques à tirer des modèles de compétences partielles et leur mise en œuvre sur le plan empirique, Beck & Klieme, 2007; Willenberg, 2007; DESI-Konsortium, 2008).

En conclusion, au vu de l'examen effectué à l'époque en vue d'élaborer une grille de compétences «langue première» pour l'interface Gymnase-Université, le modèle de DESI brièvement présenté ici s'est avéré peu propre à fournir le champ de recherche théorique requis par les orientations spécifiques qui étaient celles d'EVAMAR II.

2.2.3 PISA

2.2.3.1 Brève présentation

L'étude internationale PISA de l'Organisation de coopération et de développement économiques (OCDE) se fonde sur le concept de «*literacy*» de la critique anglo-saxonne. Celle-ci conçoit les compétences comme des instruments de base de la culture et attribue à la lecture un rôle de qualification clé, une fonction importante de prérequis à une bonne utilisation du monde des médias (Artelt, Stanat, Schneider, Schiefele & Lehmann, 2004, p. 141).

Les tests de performance couvrent un large éventail de situations de lecture axées sur la vie quotidienne et offrent aussi, à côté de textes continus, des textes appelés discontinus (liés à des tableaux, à des graphiques). Une proportion de 70% des consignes supposent des performances immanentes au texte; 30% vont au-delà, donc requièrent aussi une réflexion et une évaluation. 45% des items sont construits comme formats ouverts (Artelt et al., 2004, p. 142).

A l'inverse du modèle linguistique large du DESI, PISA se concentre sur le domaine partiel de la compréhension de l'écrit. A l'origine, le cadre conceptuel théorique du projet (OCDE, 1999/2000, 2001, cité par Artelt & Schlagmüller, 2004, p. 169) définissait cinq aspects de la lecture. Les tâches de lecture soumises par PISA devaient illustrer les capacités suivantes: premièrement, enregistrer des informations, deuxièmement, développer une compréhension générale du texte et troisièmement, une interprétation découlant du texte, quatrièmement, savoir réfléchir sur le contenu du texte et cinquièmement sur sa forme (Artelt et al., 2004, p. 143; Artelt & Schlagmüller, 2004, p. 173). Au moment même de la conception des tests, aucun modèle de compétence découlant de PISA n'était encore disponible (Artelt et al., 2004, p. 151).

Les données empiriques relevées par PISA 2000 ont permis de confirmer ensuite trois échelles des cinq qui avaient été prévues pour les compétences partielles: (1) enregistrer des informations, (2) interpréter en fonction du texte, (3) réfléchir et évaluer (Artelt & Schlagmüller, 2004, p. 175).

A partir de l'obligation de déterminer des domaines de niveaux d'étendues égales, on a formulé, pour ces trois compétences partielles de lecture, cinq degrés de compétence. Pour chaque degré, on formule une présentation des critères d'exigence caractéristiques sous forme de brefs descripteurs et de tâches exemplaires (cf. figure III.3).

Figure III.3: Le modèle de compétence de PISA pour la compréhension de l'écrit – avec cinq niveaux de compétence. - Présentation par EVAMAR des trois compétences partielles de lecture (d'après Artelt et al., 2004, p. 144-145).

PISA 2000 Compétence de lecture	Recueillir des informations	Interprétation liée au texte	Réflexion et appréciation
Niveau 5: <i>Utilisation souple de textes complexes inconnus</i>	Lecteurs expérimentés; ils comprennent des textes complexes, inconnus et longs dans leur intégralité et dans le détail, y compris des éléments ne figurant pas dans la partie principale du texte et en nette contradiction avec ce que l'on s'attend à trouver.	Le sens de fines nuances de langue est correctement interprété.	Les lecteurs expérimentés peuvent intégrer des textes longs, complexes et inconnus à leur savoir antérieur issu de divers horizons et les apprécier d'une manière critique
Niveau 4: <i>Compréhension précise de textes complexes</i>	Ces lecteurs peuvent dominer dans une large mesure les difficultés potentielles d'un texte relativement inconnu (par ex. ambiguïtés, nuances de langue ou éléments contredisant ce que l'on s'attend à trouver).	Capacité d'accéder à une compréhension détaillée de textes complexes, d'une certaine longueur et dont la forme et le contenu sont relativement inconnus.	Des textes d'une certaine longueur et assez complexes peuvent être appréciés moyennant le recours à un savoir extérieur.
Niveau 3: <i>Intégration d'éléments du texte et déduction de conclusions</i>	Les lecteurs savent utiliser des informations corollaires relativement évidentes, mais qui seraient de nature à détourner l'attention d'éléments pertinents. Ils peuvent intégrer diverses parties du texte proposé, même si l'information qui s'y trouve est peu mise en relief et que sa compréhension suppose une démarche en partie indirecte.	Niveau d'exigences moyen: capacité de construire la compréhension de textes d'un degré de difficulté moyen.	Ce qui a été lu peut être apprécié sur la base du savoir spécifique
Niveau 2: <i>Etablissement de liens simples dans le texte</i>	Les lecteurs peuvent parvenir par une simple déduction au sens de certains éléments du texte.	Capacité d'identifier l'idée principale d'un texte pas tout à fait inconnu pour construire une compréhension large du texte.	Les informations lues peuvent être appréciées en référence aux expériences et aux attitudes personnelles.
Niveau 1: <i>Compréhension superficielle de textes simples</i>	L'information nécessaire doit être clairement reconnaissable dans le texte et celui-ci ne doit comporter que peu d'éléments accessoires susceptibles de détourner l'attention de l'information centrale.	Compétences élémentaires de lecture: capacité de faire bon usage de textes simples dont la forme et le contenu sont familiers.	Seuls quelques liens relativement évidents peuvent être établis entre ce qui a été lu et le savoir commun de la vie quotidienne.

En pensant au module C d'EVAMAR II, il est intéressant de constater ici qu'il n'était pratiquement pas possible de définir, avant de passer les tests, quels pouvaient être les éléments qui allaient générer la difficulté des items des divers niveaux. (Artelt et al., 2004, p. 153).

Indépendamment des divers aspects de la lecture décrits, il convient d'analyser diverses variables: dans quelle mesure le genre et type de texte et les formats peuvent-ils influencer sur les compétences enregistrées? Pour les textes continus, il convient d'accorder une importance aux variables de la cohérence, de la densité des

propositions et du recours aux métaphores et aux analogies. Dans le cas des textes non continus, d'autres éléments générateurs de difficultés peuvent provenir des processus de traitement de l'image. On a pu prouver empiriquement, pour PISA 2000, que «la différenciation de la compétence de lecture selon qu'elle s'exerce sur des textes continus ou non continus est justifiée par les valeurs recueillies.» (Artelt & Schlagmüller, 2004, p. 176).

Cette distinction reçoit un indice de validité supplémentaire dans l'échantillonnage des résultats de PISA différenciés selon les sexes: l'avantage au score des jeunes filles sur les garçons est nettement plus petit dans la lecture de textes non continus que dans la lecture des textes continus, d'une manière statistiquement significative (Artelt & Schlagmüller, 2004, p. 176-177, 183-185).

En outre, la familiarité avec les textes littéraires peut aussi être considérée comme une dimension particulière de la compétence de lecture (Artelt & Schlagmüller, 2004, p. 178).

Si l'on tente de s'abstraire des scores, des résultats et des conclusions et de se représenter le modèle propre à servir à l'élaboration d'une grille de compétences, on pourrait utiliser les savoir-faire partiels de lecture que l'expérience a confirmés, reprendre les cinq échelons de compétences employés par PISA ainsi que les divers types et genres de textes distingués: la matrice d'un modèle de lecture aurait alors l'allure représentée dans la figure III.4).

Figure III.4: Présentation par EVAMAR: le modèle de lecture PISA avec intégration des genres de textes envisagés.

PISA 2000 Compétence de lecture	Textes non continus	Textes continus non littéraires	Textes littéraires
Recueillir des informations	Niveau 5	Niveau 5	Niveau 5
	Niveau 4	Niveau 4	Niveau 4
	Niveau 3	Niveau 3	Niveau 3
	Niveau 2	Niveau 2	Niveau 2
	Niveau 1	Niveau 1	Niveau 1
Interpréter des éléments liés au texte	Niveau 5	Niveau 5	Niveau 5
	Niveau 4	Niveau 4	Niveau 4
	Niveau 3	Niveau 3	Niveau 3
	Niveau 2	Niveau 2	Niveau 2
	Niveau 1	Niveau 1	Niveau 1
Réfléchir et évaluer	Niveau 5	Niveau 5	Niveau 5
	Niveau 4	Niveau 4	Niveau 4
	Niveau 3	Niveau 3	Niveau 3
	Niveau 2	Niveau 2	Niveau 2
	Niveau 1	Niveau 1	Niveau 1

2.2.3.2 Appréciation du modèle PISA pour une utilisation éventuelle dans EVAMAR

Le modèle PISA pourrait être aménagé en fonction des exigences d'EVAMAR II: à cette fin, il faudrait le centrer sur les niveaux supérieurs et l'enrichir d'un niveau supplémentaire 6.

Seul serait couvert le domaine partiel de la lecture. Il serait pensable d'adopter les sous-échelles «Recueillir des informations - interpréter des éléments liés au texte - réfléchir et évaluer» comme des compétences partielles de «Lire» validées empiriquement, bien qu'il faille examiner l'adaptation à la lumière des structures de processus

plus complexes. La distinction des textes non continus et continus (et la présence de la variante des textes littéraires) sont également judicieuses, à ceci près que la catégorie des textes continus non littéraires est majoritaire et que les autres catégories n'apparaîtraient qu'isolément dans l'une ou l'autre discipline (par ex. les textes littéraires en germanistique, les textes non continus en mathématiques). En bref, il y aurait lieu de transposer le profil des tests prévus pour la 9^e année scolaire aux conditions des douzième et treizième années de scolarité et aux exigences et attentes de l'université.

2.2.4 Le Cadre européen commun de référence CECR

2.2.4.1 Brève présentation

Le Cadre européen commun de référence CECR, adopté en 2001 par le Conseil de l'Europe, offre un cadre complet, facile d'emploi et cohérent pour l'apprentissage et l'enseignement des langues. Il découle d'une discussion menée dès 1971 par les experts de la didactique des langues de 40 pays et constitue la base commune qui fournit des outils de travail concrets dans la mise en œuvre de la politique des langues du Conseil de l'Europe. On recourt au CECR en fonction des objectifs pour élaborer des cursus linguistiques, des directives sur les cursus et les cours, des examens et des moyens d'enseignement. Les niveaux de compétence sont définis dans la perspective d'un apprentissage tout au long de la vie. (Conseil de l'Europe, 2001, p. 9).

Le point d'ancrage du CECR est pragmatique: les utilisateurs des langues sont considérés dans leur statut d'acteurs sociaux; il s'ensuit que les descripteurs de compétences langagières s'inscrivent dans des situations concrètes d'utilisation de la langue donnée (Conseil de l'Europe, 2001, p. 15, 18). En l'occurrence, le CECR se concentre sur les domaines essentiels, soit l'utilisation de la langue dans le monde éducatif, dans les domaines professionnel, privé et public (p. 18). Les échelles de référence elles-mêmes doivent rester libres de tout contexte explicite, c'est-à-dire être susceptibles d'être appliquées à tout contexte. Sa simplicité même fait du CECR un instrument d'un emploi aisé, et les supports électroniques et imprimés servant à l'évaluation et à l'auto-évaluation et aux tests permettent un emploi varié pour l'enseignement des langues centré sur l'élève et dans la préparation de travaux d'évaluation.

Le CECR offre un cadre permettant de préparer des critères objectifs pour décrire des compétences langagières sur une base large et solide. A ce cadre appartiennent d'abord les compétences générales des utilisateurs de la langue, leurs connaissances (ou savoir déclaratif, savoir), leurs capacités pragmatiques (savoir-faire) ou leurs compétences liées à leur personnalité propre et caractérisées par leurs attitudes (savoir-être) ainsi que leur aptitude à progresser (savoir apprendre). Deuxièmement, on considère que la compétence à communiquer «langagièrement» comporte plusieurs composantes: linguistique, sociolinguistique, et pragmatique. Les compétences linguistiques englobent les connaissances et savoir-faire lexicaux, phonologiques et syntaxiques, tandis que les compétences sociolinguistiques dépendent de la connaissance des conditions des conventions sociales. Quant aux compétences pragmatiques, elles règlent l'emploi fonctionnel des ressources de la langue et comportent – notamment – l'identification des sortes de textes et types de textes, l'emploi adéquat de l'ironie et la capacité de comprendre et de construire de la cohérence et de la cohésion¹⁰. L'ensemble des activités engageant concrètement la langue et la communication (*language activity*) distingue les domaines de la réception, de la production, de l'interaction et de la médiation, laquelle vise essentiellement l'interprétariat ou le résumé. (Conseil de l'Europe, 2001, p. 16-18).

L'espace de compétences pour les apprenants est couvert dans le CECR par un cadre de référence en six niveaux larges d'A1 à C2, auxquels on peut faire correspondre globalement la subdivision traditionnelle des niveaux débutant, moyen et supérieur. A1 est sans doute le niveau le plus élémentaire que l'on puisse identifier de

¹⁰ Le terme de *cohérence* s'applique à la structure profonde d'un texte, tandis que *cohésion* se réfère à sa structure de surface. – Quant à la *compétence textuelle*, elle est le savoir-faire d'une personne capable, quand elle lit, «de passer de la saisie des significations linguistiques des éléments de surface à la compréhension du sens global et profond du texte» («vom Verstehen der sprachlichen Bestandteile der Textoberfläche zum Sinnverständnis des Textes vorzudringen» Hornung, 2005, p. 129).

capacité à produire de la langue à titre personnel. A ce stade, on peut demander à l'apprenant et usager des tâches simples et globales (par exemple «saluer de manière simple» ou «remplir un formulaire simple avec son nom, etc.») (Conseil de l'Europe, 2001, p. 30). Le niveau A, de l'utilisateur élémentaire, est donc le degré débutant dans l'apprentissage de la langue, il est étalonné en deux échelons A1 et A2 (A1 *Breakthrough*: niveau introductif ou découverte et A2 *Waystage*: niveau intermédiaire ou de survie). Lui succède le plan B, pour l'utilisateur indépendant de la langue, où l'on distingue aussi deux degrés (B1 *Threshold*: niveau seuil et B2 *Vantage*: niveau avancé), ce dernier supposant déjà des «réponses appropriées dans des situations courantes» (Conseil de l'Europe, 2001, p. 24). Le troisième degré C concerne l'utilisateur expérimenté de la langue, il est subdivisé en niveau autonome C1 (*Effective Operational Proficiency*), qui suppose la maîtrise de tâches de communication complexes dans la profession et les études, et le niveau le plus élevé C2 (*Mastery*: maîtrise), qui correspond à l'examen le plus élevé dans l'échelle de l'Association des centres d'évaluation en langue en Europe ALTE (Association of Language Teachers in Europe, Conseil de l'Europe, 2001, p. 30-34).

Il s'agissait de décrire concrètement les savoir-faire correspondant aux six niveaux A1 à C2 du CECR; pour la Suisse, l'échelle et ses descripteurs de compétence ont été élaborés entre 1993 et 1996 à l'occasion d'un projet du Fonds national suisse s'inscrivant dans le programme national de recherche PNR 33 «Efficacité de nos systèmes de formation» (Schneider & North, 1999). Formulés à la première personne, comme des «capacités de faire» positives avec des exemples, les descripteurs ont été étalonnés en fonction du CECR suivant la méthode Rasch (Conseil de l'Europe, 2001, p. 155-160). La base d'enquête a été fournie par environ 2800 apprenants appartenant à 500 classes des degrés secondaires I et II de la formation professionnelle et de la formation des adultes. Ce projet de Schneider et North pour l'élaboration des descripteurs a permis d'introduire des niveaux intermédiaires plus fins définis empiriquement (ou appelés «niveaux plus» A2+/B1+/B2+) qui rendaient probable une échelle cohérente de 9 niveaux (Conseil de l'Europe, 2001, p. 30, 155-157) ou de 10 niveaux (Schneider & North, 1999, p. 17); dans la pratique cependant, on ne se réfère non à cette «échelle suisse» pour des raisons politiques et historiques, mais bien le plus souvent aux six niveaux de compétence du CECR décrits plus haut (Schneider & North, 1999, p. 16; Conseil de l'Europe, 2001, p. 24).

L'échelle dans son ensemble (cf. figure III.5) contient des descripteurs indiquant des «capacités de faire» générales valables pour les six niveaux; dans la grille globale servant à l'autoévaluation (p. 26) ces descripteurs se retrouvent ordonnés selon les situations d'utilisation de la langue Comprendre, Parler, Ecrire.

Figure III.5: Niveaux communs de référence: échelle globale (extrait niveau C) du CECR
(Conseil de l'Europe 2001, p. 25)

Utilisateur expérimenté	C2	Peut comprendre sans effort pratiquement tout ce qu'il/elle lit ou entend. Peut restituer faits et arguments de diverses sources écrites et orales en les résumant de façon cohérente. Peut s'exprimer spontanément, très couramment et de façon précise et peut rendre distinctes de fines nuances de sens en rapport avec des sujets complexes.
	C1	Peut comprendre une grande gamme de textes longs et exigeants, ainsi que saisir des significations implicites. Peut s'exprimer spontanément et couramment sans trop apparemment devoir chercher ses mots. Peut utiliser la langue de façon efficace et souple dans sa vie sociale, professionnelle ou académique. Peut s'exprimer sur des sujets complexes de façon claire et bien structurée et manifester son contrôle des outils d'organisation, d'articulation et de cohésion du discours.

Pour chaque domaine langagier, plusieurs sous-échelles permettent une évaluation plus différenciée.

Dans le domaine de la production, l'évaluation se fonde sur la grille concernant la communication orale (Schneider & North, 1999, p. 22; Conseil de l'Europe, 2001, p. 28) et sur cinq sous-échelles pour le monologue suivi, l'oral centré sur l'argumentation et l'exposé (Conseil de l'Europe, 2001, p. 49-50), ainsi que sur plusieurs sous-échelles pour l'interaction orale comme «s'exprimer oralement en continu» ou «participer à des conversations» (Conseil de l'Europe, 2001, p. 61-68).

Pour l'écrit, le Conseil de l'Europe (2001) présente des échelles s'appliquant à la production écrite en général, à l'écriture créative et à la rédaction de rapports et de compositions¹¹ (p. 51-52), puis à l'interaction créative (p. 68-69) ainsi qu'à «remanier des textes» et à «prendre des notes (lors d'exposés, de séminaires etc.)» (p. 77). Trois autres échelles subsidiaires existent en relation avec les stratégies générales de production avec planifier, compenser et contrôler (p. 53-54). Pour les stratégies liées à l'interprétariat (par ex. résumer, paraphraser des textes spécialisés), il n'existe pas encore de grilles.

Dans le domaine de la réception, les cinq échelles d'exemples suivantes concernent la lecture: compréhension générale de l'écrit, comprendre la correspondance, lire pour s'orienter, lire des annonces et des instructions, comprendre des directives écrites (p. 57-59).

Le domaine de la réception audio-visuelle est couvert par l'échelle «comprendre des émissions de télévision et des films» (p. 59) et celui de la compréhension de l'oral, par plusieurs sous-échelles (p. 54-56). En outre, à propos des stratégies de réception, importantes pour intégrer le contexte sur la base d'un savoir général pertinent, il existe l'échelle partielle «recueillir des indices et faire des déductions» (p. 60).

Le domaine des compétences communicatives langagières offre l'échelle globale «étendue linguistique générale» et les sous-échelles «étendue du vocabulaire» (p. 88) et «maîtrise du vocabulaire» pour le lexique (p. 89), pour la grammaire, l'échelle «correction grammaticale» (p. 90) et, s'agissant des graphies, l'échelle «maîtrise de l'orthographe» (p. 93). En outre, le champ sociolinguistique est couvert par l'échelle «correction sociolinguistique» (p. 95), et la compétence pragmatique donne lieu aux quatre sous-échelles du domaine

¹¹ Les descripteurs de ces trois sous-échelles n'ont pas été calibrés d'une manière empirique, mais créés par une nouvelle combinaison d'éléments de descripteurs extraits d'autres échelles (p. 51). Cette précision s'applique aussi aux sous-échelles de l'oral et à l'échelle de l'orthographe.

discursif «souplesse», «tour de parole», «développement thématique» et «cohérence & cohésion» (p. 97-98) ainsi que, plus techniquement, «aisance à l'oral» et «précision» (pour l'écrit, p. 100-101).

Le projet DIALANG, qui bénéficie du support financier de la Commission européenne, a mis au point des échelles descriptives supplémentaires et les a soumises à une validation empirique. Après étude de calibrage menée sur des données finlandaises suivant la méthode Rasch au centre de linguistique appliquée de l'Université de Jyväskylä (FIN), un choix de 107 items retenus a été mis à disposition pour l'auto-évaluation (Conseil de l'Europe, 2001, p. 161-172). La figure III.6 présente un extrait du niveau C.

Les spécifications de DIALANG, traduites de la version anglaise en 14 langues européennes, comprennent un test d'auto-évaluation, un test de langue et un système interactif d'aide et de conseil auquel tous les usagers peuvent avoir recours gratuitement (par le site www.dialang.org).

Figure III.6: Tableau des spécifications pour l'auto-évaluation de DIALANG – extrait du niveau C (Conseil de l'Europe, 2001, p. 165-168)

Compréhension de l'écrit	
C1	Je peux comprendre en détail un grand nombre de longs textes complexes si je peux relire les passages difficiles.
C1	Dans de longs textes complexes traitant de la vie sociale, professionnelle et universitaire, je peux identifier des finesses de signification, également des détails concernant des attitudes et opinions qui ne sont pas explicitement décrites.
C1	Je peux comprendre différents écrits en utilisant le dictionnaire à l'occasion.
C1	Je peux comprendre en détail de longues instructions complexes sur une nouvelle machine ou des opérations nouvelles même en dehors de mon domaine si je peux relire les parties difficiles.
C2	Je peux comprendre et interpréter pratiquement toutes sortes de textes écrits y compris abstraits, structurellement complexes, littéraires avec des traits familiers ou des textes documentés
C2	Je peux comprendre les subtilités de style et de sens, qu'elles soient implicites ou explicites et les différents niveaux de sens d'un grand nombre de longs textes complexes.
Production écrite	
C1	Je peux développer et défendre mes points de vue en donnant d'autres arguments, raisons ou exemples appropriés.
C1	Je peux développer méthodiquement un raisonnement en insistant sur les points significatifs et en fournissant des preuves à l'appui.
C1	Je peux donner une information claire et détaillée sur des sujets complexes.
C1	Je peux habituellement écrire sans consulter un dictionnaire.
C1	Je n'ai pas besoin de faire vérifier ou relire ce que j'écris, sauf s'il s'agit d'un texte important.
C2	Je peux structurer ce que j'écris de manière adéquate afin de faciliter le lecteur à repérer les points significatifs.
C2	Je peux écrire des rapports, des articles ou des dissertations complexes, logiques et clairs ou faire la critique de propositions ou d'œuvres littéraires.
C2	Je peux toujours écrire sans consulter un dictionnaire.
C2	Je peux suffisamment bien écrire pour ne pas avoir besoin de l'aide de quelqu'un de langue maternelle.
C2	J'écris si bien que mon texte ne puisse être amélioré de façon significative, même par un professionnel de l'écriture.

On a enregistré un autre développement dans le cadre d'un projet de recherche de longue durée des professeurs de l'ALTE (*Association of Language Testers in Europe*): une échelle de 400 items avec d'autres descripteurs «de

capacités de faire» a été mise au point pour les domaines Vie sociale et Tourisme, Travail ou Etudes; les items ont été étalonnés suivant la méthode Rasch et adaptés aux niveaux du CECR (Conseil de l'Europe, 2001, p. 173-184).

Lorsque la structuration des compétences langagières se fait selon les deux pôles de l'oral (écouter/parler) vs écrit (lire/écrire), il faut observer, pour affiner la réflexion théorique, que tant les compétences de réception que de production ne s'actualisent pas dans une seule dimension orale (du parler – écouter, écouter – parler) ou écrite (du lire – écrire, écrire – lire), mais aussi en se combinant, tant oral – écrit (écouter – écrire; parler – écrire) que écrit – oral (lire – parler; écrire – parler) (Hornung, 2005, p. 128). L'influence ou l'interdépendance¹² de compétences que ces combinaisons manifestent sont prises en compte par exemple dans l'échelle auxiliaire du CECR «prendre des notes (conférences, séminaires, etc.)» puisque cette grille globale n'est pas entièrement attribuée à la production écrite, mais traite d'une réponse psycholinguistique à un stimulus oral et touche à l'influence réciproque du texte et du média (Conseil de l'Europe, 2001, p. 75-78).

EVAMAR II a pu se permettre de négliger ces différenciations qui seraient éventuellement à prendre en compte ultérieurement, étant donné que tant la mesure des performances du module C que les analyses des examens écrits de maturité dans le sous-projet D1 impliquent des démarches liées uniquement à l'écrit (lire – écrire, c'est-à-dire réception d'un texte puis production d'un texte).

2.2.4.2 Appréciation du CECR pour une utilisation effective dans EVAMAR II

Une intégration dans la communauté européenne des langues paraît judicieuse, elle a déjà été menée assez loin avec le projet de recherche de Schneider et North (1999) visant au développement commun du Cadre européen commun de référence. Ce Cadre européen commun de référence pour les langues CECR offre un système arborescent souple permettant aux institutions qui y recourent de différencier et d'affiner au besoin les «branches» les plus pertinentes pour elles, tout en trouvant une place à l'intérieur du cadre général (Schneider & North, 1999, p. 16; Conseil de l'Europe, 2001, p. 30-32). De même, le DESI, au nombre des analyses de performances scolaires les plus importantes menées à l'échelle de l'Allemagne (cf. section 2.2.2), s'est référé au Cadre européen commun de référence CECR en tenant compte de la conception générale de ce dernier pour développer ses propres items de tests (Beck & Klieme, 2007, p. 2, 6; Nold, 2007; Takala, 2007).

A l'inverse de PISA, le CECR couvre l'intégralité des compétences langagières: lire et écouter (en tant que compréhension), parler et écrire – tout autant donc la réception et la production que l'interaction – de même que d'autres savoir-faire liés à la communication et qui peuvent être rangés dans le domaine de la conscience du fonctionnement de la langue. Tous les domaines ont reçu des différenciations par le biais de sous-échelles, définies et calibrées le plus souvent de manière empirique. La production orale y est dotée d'une pondération importante, car cette dimension est pertinente pour l'apprentissage des langues, même si elle n'est pas aussi essentielle pour les études universitaires; EVAMAR II a donc choisi de lui accorder une importance réduite.

Le projet EVAMAR II a été en mesure – sur la base de notre appréciation du CECR – de s'intégrer au profil de ce CECR pour le domaine à analyser du niveau C (utilisateur expérimenté). Quelques aménagements et un complément pour les types des textes («textes scientifiques universitaires») ont été nécessaires. En outre, il a fallu définir avec précision les domaines de compétences à analyser.

Il s'est agi aussi d'examiner l'éventualité d'ajouter un degré à ceux du CECR en C3 ou D pour l'utilisateur universitaire. Cette option a été clairement rejetée par l'équipe d'experts des partenaires du projet.

¹² Une réception réussie du message oral doit précéder la prise de notes, afin que la mise par écrit puisse avoir lieu avec succès, – que celle-ci ait la qualité d'un résumé ou d'une mise en évidence; les processus qui interviennent dans la réception et la pondération des informations doivent être considérés comme des compétences partielles de nature cognitive.

EVAMAR II a pu construire son travail sur les descripteurs de «capacité de faire» du niveau C repris du CECR, et compléter ceux-ci d'autres rubriques concernant le type des textes universitaires.

Compte tenu du fait que le CECR propose «une tentative guidée méthodique pour trouver une solution au premier problème sous-jacent relatif à l'apprentissage des langues vivantes en contexte européen», c'est-à-dire le problème de la comparabilité (Conseil de l'Europe 2001, p. 138), il s'est avéré logique de situer les analyses d'EVAMAR II dans le cadre théorique de ce système susceptible de développement et d'extension.

2.2.5 HarmoS

2.2.5.1 Brève présentation

Durant la période de mise en route d'EVAMAR II, les travaux d'un autre projet se trouvaient très avancés: à savoir le développement d'un modèle de compétence pour la langue première qui serait rendu impératif sur l'ensemble du pays par l'adhésion des cantons. Comme «l'accord intercantonal sur l'harmonisation de la scolarité obligatoire (HarmoS)» allait être la base de standards de formation impératifs pour la scolarité obligatoire, il était judicieux qu'EVAMAR II examine l'éventualité de recourir au travail d'HarmoS pour ses analyses.

L'un des objectifs de «l'accord intercantonal sur l'harmonisation de la scolarité obligatoire (HarmoS)» est de mettre au point la description des compétences et d'élaborer des standards de formation comme instruments de développement et d'assurance qualité (CDIP, 2007, art. 7); les travaux scientifiques assurent le cadre de travail du projet (CDIP, 2004a, 2004b, 2007). Une étape ultérieure prévoit la coordination des plans d'études, moyens d'enseignement, instruments d'évaluation et standards de formation (CDIP, 2007, art. 8).

Les modèles de compétences à élaborer dans le cadre de HarmoS, et qui feront l'objet d'une validation empirique, devront inclure des descripteurs de niveaux et permettre de fixer, par décision politique, des standards de prestations par discipline pour trois moments de la scolarité: la fin des deuxième, sixième et neuvième années scolaires. Les objectifs et les conceptions d'HarmoS se fondent sur la définition théorique de compétence de Weinert (1999, 2001) et sur l'expertise de Klieme, avec la détermination du modèle de compétences qui en fait partie (Klieme et al., 2003, p. 74) (EDK, 2004a, p. 8, p. 9; Sieber, 2006b).

S'agissant de l'élaboration d'un modèle de compétences HarmoS, le directeur du consortium de recherche pour la «langue première» (ou langue scolaire locale) trace la démarche à venir (Sieber, 2006a, 2006b): une présentation en forme de matrice (cf. figure III.7) donne une vue d'ensemble simplifiée de l'activité langagière, le premier plan structure la réalité en cinq aspects (situer, planifier, réaliser, évaluer, réparer; plan 1), le deuxième distingue quatre modalités (écouter, parler, lire, écrire; plan 2), Les champs d'activité ainsi définis peuvent recevoir les descripteurs de capacités (plan 3, Sieber, 2006a, p. 15).

Figure III.7: HarmoS: Un modèle pour la description (Sieber, 2006a, p. 18)

		Domaines partiels de l'activité langagière (plan 2)			
		Ecouter	Parler	Lire	Ecrire
Aspects d'activités langagières (plan 1)	Situer	Tâches (plan 3)	Tâches	Tâches	Tâches
	Planifier	Tâches	Tâches	Tâches	Tâches
	Réaliser	Tâches	Tâches	Tâches	Tâches
	Réparer	Tâches	Tâches	Tâches	Tâches
	Evaluer	Tâches	Tâches	Tâches	Tâches

L'articulation détaillée des tâches et des compétences partielles pour chacun des domaines d'activités langagières est illustrée par l'exemple de l'écoute dans la figure III.8

Figure III.8: HarmoS – un exemple: savoir écouter (Sieber, 2006a, p. 19)

Activité langagière	Ecouter
Situer	Reconnaître la fonction de communication Identifier la forme de communication Identifier les styles de communication (...) Garder à l'esprit le but de la communication Relier ce qu'on entend à ce qu'on sait Activer et maintenir la motivation et l'intérêt
Planifier	Identifier le plan du discours Sélectionner les stratégies d'écoute Structurer les aspects linguistiques et le savoir non-linguistique
Réaliser	Activer les savoirs (linguistiques et généraux) (...) Se concentrer sur le message du locuteur Respecter les règles de l'écoute Comprendre les messages verbaux (...) / non verbaux, établir la cohérence, construire le sens Utiliser les moyens auxiliaires (notes, tableau...)
Réparer	Réécouter ce qu'on n'a pas compris quand c'est possible (enregistrement) (...) Récapituler ce qu'on a compris du texte
Evaluer	Exercer un contrôle permanent de l'écoute Estimer l'effet de la parole d'autrui Contrôler la construction du sens globale et de détail, procéder aux adaptations Poser des questions de compréhension, aller au-delà, demander des explications Réfléchir au contenu et à la forme

A propos de la compétence de lecture (compréhension de l'écrit), HarmoS tient compte de la variété des savoir-faire en distinguant d'une part les formats de textes «continus vs non continus», à l'instar de PISA (cf. section 2.2.3) et d'autre part les genres de textes et leurs diverses fonctions. HarmoS assemble les genres de textes en trois regroupements: celui des textes narratifs (littéraires), celui des textes descriptifs ou explicatifs, et celui des textes argumentatifs (Sieber, 2006c, p. 24-25).

2.2.5.2 Appréciation du modèle HarmoS pour son utilisation dans EVAMAR II

La discussion à propos du projet HarmoS a essentiellement porté sur l'utilisation éventuelle du modèle de compétences qui s'y trouve comme d'une variante à l'insertion de la grille de compétences d'EVAMAR II dans le CECR (cf. section 2.2.4). Or, il n'a pas été possible de reprendre directement le modèle HarmoS à la suite des considérations suivantes: Les paramètres pertinents pour EVAMAR II, tels qu'ils sont esquissés plus haut (III,1), sont les années concernées (tardives), le type gymnasial pour l'école envisagée et le degré élevé des compétences attendues. On pouvait dès lors douter sérieusement de la possibilité d'aligner les tests d'EVAMAR sur les trois moments de mesure retenus par HarmoS (soient les termes des 2^e, 6^e et 9^e années de la scolarité). En outre, le type des textes scientifiques universitaires spécialisés, qui devaient servir de base aux tests de performances EVAMAR en langue première, a été un autre critère important qui rendait difficile le recours au projet HarmoS.

Vu l'importance de HarmoS pour la période pré-gymnasiale, on a envisagé la possibilité d'une synthèse de ce modèle issu de l'école obligatoire avec le CECR «tertiaire». En outre, nos partenaires de la Suisse romande et du Tessin nous ont suggéré de développer une variante au modèle langue première visant à combiner, dans une matrice originale, les domaines partiels de compétence du CECR avec les activités langagières retenues par HarmoS. Nous avons tenté d'analyser les descripteurs de capacités des deux systèmes et de les intégrer dans un nouveau modèle qui soit cohérent. Il est ressorti de cette démarche que les exemples projetés n'étaient guère utilisables ou même inutilisables, parce que chaque système avait ses critères de catégorisation et sa logique propre qui ne pouvaient dès lors donner lieu à des cellules aisées à distinguer. Nous avons donc abandonné cette voie.

2.2.6 Relevé de l'état des prestations scolaires dans le land de Rhénanie-du-Nord-Westphalie (NRW)

2.2.6.1 Brève présentation

Une autre possibilité de référence se présentait à l'équipe d'EVAMAR II, dans son effort d'élaborer une grille de compétences pour la langue première: le modèle qui sert de base au relevé annuel des prestations scolaires en Rhénanie-du-Nord-Westphalie (NRW): depuis l'automne 2004 en effet, tous les élèves de neuvième année (dès 2006/2007: de huitième année), au nombre de 200'000 environ, sont soumis à un relevé de leurs prestations à l'échelle du land sous l'égide du ministère compétent (*Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen*, MSWNRW). Toutes les classes sont concernées – écoles de base, globales, «Realschulen» et gymnases.

Comme le soulignent Leutner, Wirth et Fleischer (2005), ces relevés ne se limitent pas à enregistrer des prestations scolaires au sens de performance (*Performanz*) selon Prenzel, Drechsel, Carstensen et Ramm (2004), mais ils envisagent bien plus les compétences dans la définition que Weinert (2001) donne du concept quand il parle de «savoir-faire potentiels» (potentialités d'attitudes «*Fähigkeitspotenziale*») ou, pour prendre le langage de Klieme et al. (2003, p. 72), au sens de «disposition qui rend les élèves capables de» («*befähigende Disposition*», p. 13).

Les modèles de compétences spécifiques aux disciplines qui ont présidé à l'élaboration des données soumises aux élèves permettent de ranger les items à des niveaux de compétence déterminés, avec divers degrés de difficultés. Dans la suite de notre compte rendu, nous donnons les niveaux de compétence tels qu'ils sont définis pour l'allemand dans les domaines testés du relevé des prestations scolaires du land (MSWNRW, 2005b).

Pour le premier domaine – «savoir lire» compréhension de l'écrit/réflexion sur la langue (abrégié D-LV pour *Leseverstehen/Reflexion über Sprache*), on a distingué cinq niveaux de compétence: le premier niveau (D-LV 1) est décrit comme la simple compréhension de l'information (MSWNRW, 2005b) ou la compréhension de l'écrit dans ses débuts (MSWNRW, 2007), le deuxième niveau (D-LV 2) concerne la capacité de compréhension simple de l'écrit, le troisième niveau (D-LV 3), la compréhension approfondie de l'écrit, le quatrième niveau (D-LV 4) est celui de la compréhension différenciée de l'écrit et enfin le cinquième niveau (D-LV 5) englobe la compréhension de l'écrit complexe (MSWNRW, 2005b, p. 1-3; MSWNRW, 2007).

Le deuxième domaine, «écouter et intégrer», concernant l'écoute et l'assimilation (D-ZV pour *Zuhören und Verarbeiten*), est subdivisé en trois niveaux avec un niveau supplémentaire: Le premier niveau (D-ZV 1) correspond à la prestation de «saisir et rendre des informations simples transmises par oral, au deuxième niveau (D-ZV 2), l'élève est capable de saisir, identifier et ordonner des informations transmises de façon claire par oral, au niveau de compétence trois (D-ZB 3), l'élève peut saisir des informations d'une manière différenciée et effectuer des opérations d'assimilation complexes, au niveau D-ZV 3+ enfin, une personne est capable de saisir sans aide des informations de façon différenciée et de les mettre en ordre de manière autonome (MSWNRW, 2005b, p. 4-5).

Pour le troisième domaine («écrire», abrégé D-S pour *Schreiben*), le premier degré de compétence (D-S 1) vise la simple saisie et présentation d'un contenu ou d'un sujet, le deuxième niveau (D-S 2) la présentation ciblée d'un contenu en fonction d'un objectif ou d'une tâche, et au niveau 3 le plus élevé (D-S 3), les élèves ont la capacité de suivre de façon organisée et en fonction d'un objectif une démarche de rédaction complexe comprenant des appréciations personnelles et des explications (MSWNRW, 2005b, p. 5-6).

EVAMAR II ne pouvait adopter que les niveaux de compétence supérieurs. La figure III.9 présente ce que donnerait une telle récapitulation. La matrice contient des descripteurs de capacités: une colonne particulière est réservée au domaine partiel de la réflexion sur la langue, que les relevés NRW intègrent à la compréhension de l'écrit.

2.2.6.2 Appréciation du relevé des prestations scolaires en NRW pour son utilisation dans EVAMAR II

Sur la base des observations et critères évoqués à la section 2.1 – accent mis sur les années «tardives» pour EVAMAR II, et sur les années gymnasiales – alors que les relevés du land NRW englobent tous les types d'école, on a envisagé de se limiter au niveau supérieur des compétences modélisées pour les domaines de l'allemand testés en NRW. Il est ressorti d'un examen plus détaillé que les trois domaines de l'allemand retenus par l'instruction publique de NRW («Leseverstehen/Reflexion über Sprache», «Zuhören und Verarbeiten» et «Schreiben» MSWNRW) n'étaient pas appropriés pour la construction des tests d'EVAMAR II à cause des restrictions qu'ils comportent: d'abord, EVAMAR II devait considérer que les domaines «écrire» et «écouter» ne pouvaient faire l'objet d'une évaluation complète dans le temps dévolu aux tests, et que des corrections du modèle auraient pris trop de temps. Ensuite, un autre problème se posait: les données élaborées année après année par le NRW sur la base de standards internes et, de ce fait, l'impossibilité de garantir à long terme la comparabilité des travaux (MSWNRW, 2005b, p. 1) rendaient difficile, pour ne pas dire impossible, l'élaboration d'un modèle généralisable.

Figure III.9: Les niveaux supérieurs de compétences dans les relevés des apprentissages en NRW 2005 – matrice refaite

Domaines testés NRW 9	Comprendre l'écrit/réfléchir sur la langue		Ecouter et intégrer	Ecrire
Niveaux de compétence	D-LV 5: Compréhension de l'écrit complexe La compétence de lecture implique des démarches intellectuelles complexes d'intégration. Outre la saisie du texte dans son ensemble, le contenu et la structure font aussi l'objet d'une observation et d'une explication précise. Les marqueurs du texte et leurs valeurs peuvent être expliqués avec leurs nuances.		D-ZV 3 plus: Saisir sans aide et de façon nuancée des informations reçues par le canal oral et en faire un usage autonome	D-S 3: Conduire clairement et en fonction d'un but une démarche complexe de rédaction en y incluant des explications et des jugements personnels
Compétences partielles décrites	Dégager et relier des informations, même quand elles ne sont pas faciles à trouver	Connaître les notions fondamentales de la grammaire et les appliquer dans des contextes donnés	Savoir tirer - sans aide - des conclusions d'une contribution radiophonique entendue et les présenter clairement	Organiser des contenus, des idées et des messages - en fonction de la consigne reçue - sous plusieurs aspects et d'une manière claire et compréhensible
	Comprendre dans le détail les diverses parties du texte et leurs relations	Donner une appréciation et une explication nuancée sur le contenu et la structure de phénomènes de langue	Savoir tenir compte aussi d'aspects abordés brièvement pour les intégrer à la tâche demandée	Introduire divers points de vue, les expliquer et les évaluer
	Reconnaître et expliquer des ambiguïtés dans les textes	Faire appel à un savoir et à des concepts spécifiques déterminés pour comprendre et expliquer le texte	Savoir présenter le contenu du texte d'une manière qui corresponde à la situation	Adapter ses propres idées et représentations aux divers aspects de la consigne reçue
	Assurer une élaboration correcte de la signification globale du texte	Réfléchir sur le contenu, la langue et la structure du texte dans son ensemble		Faire intervenir des stratégies d'écriture de façon ciblée et adaptée à la consigne reçue
	Situer une affirmation donnée d'un texte complexe dans un contexte général	Evaluer des possibilités de sens en les opposant à d'autres hypothèses proches.		
	Discuter diverses hypothèses de signification à l'aide d'éléments précis du texte	Intégrer un savoir utilisé dans la vie quotidienne ou tiré du savoir général (extra-contextuel) pour élaborer la compréhension du texte ou son évaluation critique		
	Saisir le message global d'un texte littéraire complexe ou d'un texte spécialisé			

2.3 Modèle final de la grille de compétences «langue première» pour EVAMAR II

La situation était claire: le modèle final se fonderait sur des composantes à choisir parmi les modèles de compétences décrits à la section 2.2, en tenant compte des aspects et critères exposés dans les sections 1 et 2.1. Les appréciations de détail se trouvent à chaque fois dans la section consacrée aux études qui y sont traitées. Certaines parties du CECR présentent la meilleure correspondance avec nos critères (cf. appréciation dans la section 2.2.4.2). Le paragraphe qui suit comporte une explication plus approfondie sur le CECR.

2.3.1 Motivation du choix du CECR comme base de travail

Le cadre de référence retenu, inspiré du Cadre européen commun de référence (CECR), est fort bien accepté par des centaines de milliers d'apprenants et d'enseignants des langues tant dans un apprentissage qui dure toute une vie que parmi les scientifiques et les didacticiens de l'Europe entière. Ses sous-échelles sont calibrées empiriquement et sans cesse mises à jour, vérifiées et affinées. Ouvert à tous les domaines partiels des compétences linguistiques et susceptible d'évoluer à la lumière de l'expérience acquise, le CECR se distingue clairement parmi tous les modèles linguistiques analysés. Il s'est donc avéré le mieux adapté pour décrire le début d'un apprentissage à un niveau supérieur, dans la ligne du projet EVAMAR II.

Par ailleurs, les théoriciens de l'enseignement des langues, et en particulier de la langue première, n'analysent pas les grands domaines classiques d'apprentissage et de tâches de la même manière: selon la conception didactique et l'époque (cf. la vue d'ensemble qu'en donne par exemple Boueke¹³, 1979), les champs d'apprentissage qu'ils distinguent seront au nombre de trois (par exemple chez Ingendahl¹⁴, 1999) ou de cinq (voir par exemple Glinz¹⁵, Glinz & Ramseier, 1997; Abraham¹⁶, Beisbart, Koß & Marenbach, 2003, p. 9); vu cette diversité de positions des didacticiens (Ossner, 2006, p. 7)¹⁷, il apparaît judicieux d'adopter la perspective du CECR, bien acceptée dans les pays européens et susceptible d'harmoniser notre travail en fonction des besoins qui sont les nôtres.

Nous nous sommes dès lors employés à adapter le cadre commun de référence pour les langues du CECR aux nécessités du projet EVAMAR II, en l'enrichissant, par exemple, de la catégorie des textes scientifiques spécialisés au niveau universitaire. Les descriptions de détail à ce propos figurent à la section 2.3.2.

¹³ Un résumé de Boueke (1979) rappelle l'évolution de l'enseignement de l'allemand: la discussion scientifique des années 60 à ce sujet y apparaît à travers des rapports de recherche de 1965 à 1977.

¹⁴ Ingendahl (1999) distingue, «pour structurer la théorie de la didactique, et rendre compte de l'habitude des plans d'étude et des manuels scolaires» les trois domaines de tâches de la langue maternelle que représentent 1. enseigner à parler et à écrire, 2. enseigner à comprendre les opinions et les textes, 3. enseigner à réfléchir sur la langue. Le troisième domaine, jadis appelé simplement enseignement de la grammaire, est dénommé «officiellement ... depuis plusieurs décennies réflexion sur la langue et son utilisation» (p. 1). Ingendahl ne tient pas compte de la différenciation, toujours plus marquée, entre l'oral et l'écrit, sans quoi les trois domaines cités seraient au nombre de cinq.

¹⁵ Glinz, Glinz & Ramseier (1997) adoptent les cinq domaines «Voir, écouter, parler, lire, écrire» sur la couverture de leur manuel de base «Das Schweizer Sprachbuch», mais ils articulent le contenu selon les cinq champs d'apprentissage des langues: parler et agir, lire et comprendre, écrire des textes, orthographe, grammaire – avec l'écoute intégrée dans la rubrique de parler et agir et l'étude de la langue subdivisée.

¹⁶ Abraham et al. (2003) décrivent l'état actuel de la discussion sur l'enseignement de l'allemand pour les cinq champs de travail ou de tâches traditionnels (utilisation orale, utilisation écrite de la langue, étude de la langue, orthographe) ainsi que 62 activités retenues (de lecture d'auteurs jusqu'à résumé) à l'intérieur de ces domaines.

¹⁷ La contribution la plus récente de Ossner (2006), qui recherchait un consensus avec son modèle de compétence, a permis d'approfondir la discussion, mais non de trouver un accord. Une étude détaillée des diverses écoles et didactiques et des pratiques actuellement en vigueur dans les länder allemands sortirait du cadre de notre étude.

Une combinaison de cette grille avec les modèles de compétence issus de HarmoS et de PISA a été envisagée, mais finalement abandonnée pour diverses raisons (cf. section 2.2.5.2). La variante qui eût consisté à intégrer les cellules du domaine partiel de la «Lecture» selon PISA 2000 par le biais des trois échelles de rapport élaborées empiriquement «Recueillir des informations», «Interpréter des éléments liés au texte» et «Réfléchir et évaluer» a été abandonnée sur décision de l'équipe de projet.

2.3.2 Présentation de la grille de compétences retenue par EVAMAR pour la langue première

Le modèle général projeté pour EVAMAR II s'est constitué à partir du CECR, dont il a repris les échelles partielles les plus importantes: en d'autres termes celles qui sont pertinentes pour les premiers semestres des études universitaires et qui requièrent des capacités et des savoir-faire variés. On a renoncé, notamment, aux échelles proposées pour «l'écriture créative» ou «comprendre la correspondance».

L'axe horizontal de notre matrice reprend dans ses grandes lignes l'étalonnage «européen» des niveaux du CECR, cependant uniquement le champ des compétences pertinent pour les bacheliers et les étudiants, à savoir celui de l'utilisateur expérimenté de la langue (niveau C) ainsi que la partie supérieure de l'utilisateur indépendant (niveau B). On a retenu également le niveau intermédiaire B 2.1, qui est le niveau minimal de compétence en allemand langue étrangère (DaF) requis par les universités de Suisse alémanique pour l'admission des candidats qui ont une langue maternelle autre que l'allemand et attestent par là de leur aptitude aux études. Le niveau C 1.1 est considéré comme le maximum pouvant être atteint au test d'entrée par ces mêmes étudiants en allemand «langue étrangère». Le niveau d'exigence attendu pour les étudiants ayant l'allemand comme langue maternelle est le niveau C. La proposition d'ajouter un niveau supplémentaire C 3 ou D pour l'utilisateur expérimenté de niveau universitaire ou scientifique a été clairement rejetée par l'équipe des experts des partenaires au projet. On conserve ainsi six degrés de niveaux de B 2.1 à C 2.2.

Le modèle de base pour la langue première donne donc lieu à une matrice bidimensionnelle. L'axe vertical recense les domaines de l'utilisation et d'apprentissage de la langue que sont Lire, Ecouter, Ecrire, Parler et «conscience du fonctionnement de la langue» (avec vocabulaire, grammaire et orthographe respectivement). Sur l'axe horizontal, correspondant aux degrés gymnasial et universitaire, on trouve les trois à six degrés de l'échelle de compétences de B 2.1 à C 2.2, ce qui donne lieu à 84 cellules prêtes à recevoir les descripteurs (cf. figure III.10).

Figure III.10: Modèle de base en langue première sur la base du CECR (les surfaces grisées correspondent aux domaines analysés par EVAMAR II)

			Utilisateur indépendant		Utilisateur expérimenté			
			B 2		C 1		C 2	
			B 2.1	B 2.2	C 1.1	C 1.2	C 2.1	C 2.2
			Admission à l'université pour étudiants étrangers		Maximum au test d'admission pour étudiants étrangers			
			EVAMAR II					
COMPREHENSION	de l'oral							
	de l'écrit	Compréhension générale de l'écrit	EVAMAR II					
		Lire pour s'informer						
		Comprendre des informations et des arguments						
		Comprendre des instructions écrites						
RELEXION SUR LA LANGUE	Compétence grammaticale							
	Vocabulaire	Etendue lexicale						
		Maîtrise lexicale						
	Orthographe							
PRODUCTION	écrite	Ecrire des rapports et des compositions						
		Production écrite en général						
		Interaction écrite en général						
		Travailler sur des textes						
	orale	Faire un exposé						
		Participer à des conversations						

Il était clair dès le début qu'au vu des ressources disponibles et des difficultés de mesure que cela aurait supposé, EVAMAR II n'allait pas pouvoir effectuer une analyse empirique de toutes les compétences partielles nécessaires à la réussite des premiers semestres universitaires; nous avons donc choisi le champ des compétences linguistiques et concentré notre examen sur les savoir-faire partiels liés à la lecture. En effet la compétence centrale du mode réceptif qu'est la compréhension de l'écrit accapare le plus de temps, elle constitue la base qui permet d'acquérir de nouvelles connaissances. Aucun apprentissage n'est possible, aucune appropriation d'un nouveau savoir ne peut se faire si les étudiants ne sont pas en mesure de lire et de comprendre correctement le détail des dossiers, résumés et documents de cours distribués ou mis à disposition par voie électronique, d'en saisir de façon sélective les raisonnements scientifiques souvent complexes pour intégrer les contributions de la recherche dans leurs textes et leurs argumentations.

D'une importance comparable est la compétence de l'écoute, qui joue un rôle non négligeable dans les cours et les séminaires bien entendu. Comme la lecture, la compréhension de l'oral appartient au mode réceptif des compétences linguistiques, mais elle emprunte un autre canal et se trouve souvent accompagnée de supports visuels. En raison de son cadre général, EVAMAR II a dû abandonner l'examen de cette compétence d'écoute, car il eût été très coûteux de produire les supports d'analyse – en l'occurrence des enregistrements vidéo simulés ou authentiques de cours universitaires comme bases d'items de tests standardisés –. Il est cependant souhaitable qu'une étude ultérieure puisse analyser aussi ce domaine au niveau universitaire.

L'idée de tester les activités liées à l'écrit – au sens de la rédaction de travaux spécialisés – n'a pas été, elle non plus, suivie par le module C, non seulement en raison des procédures complexes de correction, mais aussi à cause des contenus: c'est en effet seulement au cours des études que les étudiants doivent produire des travaux de proséminaire ou de séminaire – donc des textes de longueur moyenne; ce savoir-faire ne s'inscrivait pas dans l'espace-temps embrassé par les analyses d'EVAMAR II.

Le champ des recherches n'a pas pu inclure les compétences partielles de l'expression orale, l'attitude en situation de dialogue et d'entretien, ni la capacité de faire et de présenter un exposé, donc la parole en monologue et en dialogue. Le modèle EVAMAR II n'a pas tenu compte non plus du domaine de l'interaction (aspect de l'utilisation de la langue largement traité dans le CECR) ni, dans la compétence de la production orale, des deux compétences partielles de la «participation à une conversation» et du «monologue suivi».

Il était ressorti du sous-projet A4 que de nombreux professeurs se plaignaient des faiblesses croissantes des étudiants en orthographe (cf. chapitre II, section 3.3): on a donc décidé de tester les compétences d'orthographe et de grammaire comme deux domaines séparés.

Pour résumer, la grille de compétences pour la langue première se limite, dans EVAMAR II, aux domaines de la compréhension de l'écrit et de la conscience du fonctionnement de la langue. Cette limitation correspond aussi à l'orientation donnée au sous-projet A1 privilégiant la compréhension de l'écrit, c'est-à-dire l'analyse de matériaux et de textes les plus divers liés aux études universitaires.

Dans le domaine de la compréhension de l'écrit, des considérations de structure liées aux tests ont amené à opérer une fusion entre les compétences partielles de «conception générale de l'écrit» et de «lire pour s'informer», difficiles à différencier; on les a donc confondues en «compréhension générale de l'écrit». En regard, la compétence de «comprendre des informations et des arguments» a été spécifiée en «compréhension détaillée de l'écrit»¹⁸. Pour le domaine de la conscience du «fonctionnement de la langue», on a regroupé sous «vocabulaire» les deux compétences partielles de «l'étendue lexicale» et de la «maîtrise lexicale», car l'élaboration des tests a montré déjà qu'il était difficile de distinguer les deux catégories. Ce deuxième accent mis sur l'emploi du vocabulaire et le recours aisé au lexique universitaire correspond ainsi aux questions du projet A1, dans lequel la compréhension et le codage d'unités de signification du vocabulaire apparaissent comme des prérequis

¹⁸ Dans la préparation des tests, nous avons renoncé à la rubrique «Comprendre des informations et des arguments». Cette compétence partielle est analogue aux items de PISA bien connus comme «comprendre des modes d'emploi», mais elle est rarement indiquée comme consigne dans le cadre des premiers semestres.

(de connaissances et de compétences). Le troisième point de cristallisation, testé uniquement pour la langue première Allemand (pour les motifs, cf. chapitre IV du rapport) englobe les compétences grammaticales et l'orthographe. L'examen de ce dernier domaine a trouvé place dans la partie 10 du test d'allemand par un format ouvert court qui comporte aussi des éléments de la production de textes libres. Sous ce dernier point se trouvaient réunis les deux domaines de compétence de la réflexion sur la langue et de la production.

Il résulte de ce travail une grille de compétences pour la langue première (cf. figure III.11) qui adopte en axe horizontal la typologie de textes qu'EVAMAR II a déduite de son enquête: la catégorie des textes scientifiques de niveau universitaire. Ainsi, le modèle Langue première servant de grille de référence et d'interprétation théorique pour la partie C consacrée aux tests comporte encore cinq cellules qui doivent recevoir un contenu: il s'agit de descripteurs de savoir-faire, en l'occurrence des résumés et condensés des échelles d'exemples du CECR pour le niveau C2; on a laissé de côté les domaines non pertinents pour l'utilisation de la langue dans le cadre universitaire (donc pour EVAMAR II), notamment ceux des textes rédigés en langue familière et des textes littéraires¹⁹. On a introduit des formulations nouvelles – en évidence *en italique* – pour la catégorie des textes spécialisés scientifiques à laquelle EVAMAR II s'intéressait en priorité.

2.4 Synthèse

Au terme de l'analyse de contenu, nous avons une image d'ensemble peu homogène des études existantes portant sur les compétences langagières et leur mise en perspective théorique en modèles linguistiques. D'un côté, l'étude de didactique scolaire DESI, qui compare les performances scolaires des jeunes en allemand et en anglais (en 9^e année), travaille sur six domaines de compétence comportant de trois à cinq niveaux de performance. A l'inverse, le modèle pour la langue première dans PISA, pour les jeunes de 15 à 16 ans, ne porte que sur le domaine de la lecture, de la compréhension de l'écrit. En revanche, l'une des rares études portant sur l'interface Gymnase-Université était la grille d'analyse textuelle de Zurich, qui pourtant n'avait analysé que la production écrite (avec des compositions de maturité). Sur la base de cette analyse comparée des modèles de compétences linguistiques existant dans notre pays et à l'étranger, le choix d'EVAMAR II s'est finalement porté sur une grille de compétences «langue première» fondée essentiellement sur le Cadre européen commun de référence CECR (Conseil de l'Europe, 2001). Avec cette option, le projet suisse EVAMAR II se situe dans le même cadre que l'étude allemande DESI. Envisagée dans un premier temps, l'idée d'un recours à des composantes issues des modèles de compétence de HarMoS et de PISA a été finalement abandonnée.

¹⁹ Comme on l'a mentionné, la part des textes littéraires proprement dits que les étudiants rencontrent durant leurs premiers semestre, même en germanistique, est étonnamment limitée au regard des introductions théoriques; il a donc été possible de les laisser de côté.

Figure III.11: Grille de compétences en vue des tests en langue première (avec descripteurs de «capacité de faire»)

			Classement des types de textes-
			Documents universitaires
COMPREHENSION	de l'écrit (LECTURE)	<p><Compréhension générale de l'écrit></p> <p><Lire pour s'informer></p>	<p>Peut comprendre et interpréter de façon critique <presque> toute forme d'écrit (textes littéraires ou non, abstraits et structurellement complexes)</p> <p>Peut comprendre une grande gamme de textes longs et exigeants, <i>y compris des textes scientifiques</i>, et en saisir les fines nuances stylistiques et les significations implicites.</p> <p>Peut parcourir rapidement des textes longs et complexes, <i>y compris des textes scientifiques</i>.</p> <p>Peut identifier rapidement le contenu et la pertinence d'articles et de textes dans une gamme étendue de sujets scientifiques afin de décider si une étude plus approfondie vaut la peine.</p>
		<p>Compréhension détaillée de l'écrit:</p> <p><Comprendre des informations et des arguments></p>	<p>Peut comprendre une gamme étendue de textes longs et complexes <i>que les étudiants rencontrent en première année d'université</i> et en apprécier de subtiles distinctions, et aussi les points de vue ou avis implicites autant qu'explicites qui s'y trouvent.</p> <p>Peut localiser une information importante recherchée²⁰.</p>
FONCTIONNEMENT DE LA LANGUE	Vocabulaire		<p>A la maîtrise d'un vocabulaire très étendu, accompagnée de la conscience des connotations.</p> <p>Utilise le lexique toujours correctement et de façon appropriée.</p>
	Compétences grammaticales		<p>Manifeste une excellente maîtrise de la grammaire, même dans l'utilisation de moyens d'expression complexes.</p>
	Orthographe		<p>Rédige des textes exempts de fautes d'orthographe.</p>

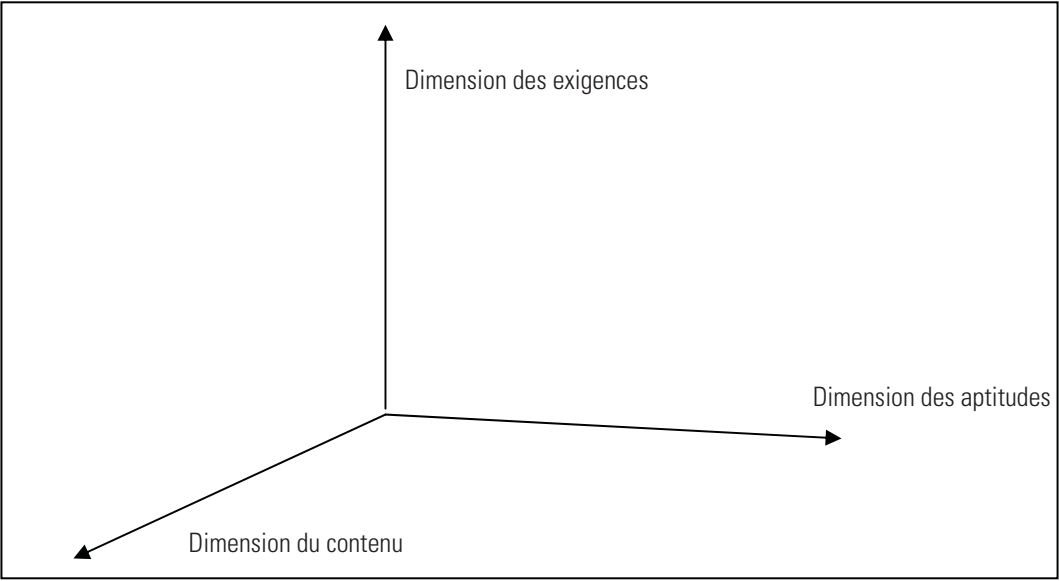
²⁰ EVAMAR II a fait passer la compétence partielle «localiser des informations importantes» du domaine de la compréhension générale de l'écrit à celui de la compréhension détaillée.

3 La grille de compétences en mathématique

3.1 Bases

L'élaboration du cadre de compétences en mathématiques se fonde elle aussi sur des bases théoriques ressortant de précédents travaux réalisés aux niveaux international, européen, suisse et régional. Le champ pertinent comprenait des études comme PISA, TIMSS ou HarmoS, dans lesquelles les mathématiques font partie des disciplines étudiées. La construction des batteries de tests correspondants en mathématiques se fonde dans la plupart des cas sur un modèle de compétences. Les modèles de compétences qui ont ainsi été développés pour les mathématiques consistent dans des schémas bi- ou pluridimensionnels décrivant les contenus, les aptitudes et éventuellement les domaines d'exigence ou les niveaux de compétence (voir figure III.12).

Figure III.12: Schéma d'un modèle de compétences à plusieurs dimensions



Les dimensions du contenu et des aptitudes sont souvent décrites sous la forme d'une matrice (voir figure III.13).

Figure III.13: Représentation en matrice des dimensions du contenu et des aptitudes

		Dimension du contenu				
		Contenu 1	Contenu 2	Contenu 3	...	Contenu n
Dimension des aptitudes	Aptitude 1					
	Aptitude 2					
	Aptitude 3					
	...					
	Aptitude m					

Les sections ci-après passent en revue quelques modèles de compétences connus pour les mathématiques en évaluant leur adéquation pour EVAMAR II. Ils se distinguent principalement par leur degré de précision et d'exhaustivité dans la description des contenus et/ou des aptitudes et dans les spécifications de la matrice.

3.2 Grilles de modèles de compétences existants pour les mathématiques

3.2.1 Le modèle de compétences du NCTM (National Council of Teachers of Mathematics)

Développé aux Etats-Unis, le modèle de compétences du NCTM (Conseil national des professeurs de mathématiques) distingue cinq paramètres de contenu (*Number & Operations, Algebra, Geometry, Measurement et Data Analysis & Probability*) et cinq critères d'aptitudes (*Problem Solving, Reasoning & Proof, Communication, Connections et Representation*) (NCTM, 2000). Ce modèle a inspiré divers autres modèles de compétences.

3.2.2 Le modèle de compétences de la KMK en Allemagne

En Allemagne, des modèles de compétences ont été développé notamment dans le cadre de la définition de standards de formation, dans un premier temps pour la 9^e et la 10^e année scolaire (voir Klieme et al., 2000).

Le modèle matriciel de la Conférence permanente des ministres de l'éducation et des affaires culturelles des länder allemands KMK (*Kultusministerkonferenz*) englobe les deux dimensions des contenus (idées-force) et des compétences mathématiques générales (voir figure III.14).

Figure III.14: Le modèle de compétences allemand de la KMK

		Idées-force				
		Nombres	Mesures	Espace et forme	Relations fonctionnelles	Données et probabilités
Compétences mathématiques générales	Argumentation mathématique					
	Résolution mathématique de problèmes					
	Modélisation mathématique					
	Utilisation de représentations mathématiques					
	Maîtrise des éléments symboliques, formels et techniques des mathématiques					
	Communication mathématique					

Une idée-force regroupe les contenus de différents chapitres mathématiques et traverse tout le cursus de mathématiques en décrivant une spirale. Les compétences mathématiques générales sont décrites précisément en fonction de leur contenu structuré par idées-force. Les exemples ci-après illustrent différentes formes de la première idée-force (les nombres) sans préciser à quelles compétences mathématiques générales elles se réfèrent:

Les élèves

- utilisent des représentations pertinentes des nombres rationnels, en particulier des nombres naturels, des nombres entiers et des fractions selon les besoins,
- représentent les nombres sous une forme adaptée au contexte, notamment en puissances de dix,
- expliquent la nécessité de l'extension progressive des ensembles de nombres en se servant d'exemples,
- utilisent les propriétés des opérations mathématiques, notamment pour faciliter leurs calculs,

- contrôlent la plausibilité de leurs calculs par des calculs approximatifs et d'autres procédés,
- arrondissent les résultats de leurs calculs de façon pertinente selon le contexte,
- manient correctement le calcul des pourcentages et des intérêts,
- illustrent les relations entre les opérations de calcul par des exemples et se servent de ces relations,
- choisissent, décrivent et évaluent les approches et les procédés qui sont à la base d'algorithmes ou d'équations,
- raisonnent en termes de combinaisons pour déterminer le nombre de possibilités dans des situations concrètes,
- examinent et interprètent des résultats présentés dans un contexte scientifique en se fondant sur une appréciation critique du modèle choisi et de sa mise en œuvre.

La résolution de problèmes mathématiques requiert des compétences mathématiques générales plus ou moins développées. On distingue trois classes de compétences:

- Compétences de classe I – Reproduction: capacité de restituer et d'appliquer directement des concepts, des propositions et des procédés de base dans un domaine déterminé et dans un cadre répétitif.
- Compétences de classe II – Mise en relation: capacité de traiter des objets connus en reliant des connaissances, des capacités et des aptitudes acquises par l'exercice dans différents domaines des mathématiques.
- Compétences de classe III – Généralisation et compréhension en profondeur: capacité de traiter des objets complexes notamment dans le but de formuler des questionnements, des solutions, des démonstrations, des conclusions, des interprétations ou des appréciations originales.

De façon générale, le niveau d'exigence et la complexité cognitive des classes de compétences augmentent dans l'ordre indiqué.

3.2.3 Le modèle HarmoS de compétences en mathématiques

Les explications ci-après sont tirées du rapport intermédiaire du projet HarmoS (Linneweber & Wälti, 2006).

HarmoS utilise un modèle matriciel qui distingue entre composantes par rapport à des contenus et compétences par rapport à des processus. A la différence d'autres modèles, le modèle HarmoS met ces deux composantes explicitement en relation. En effet, ce dont les élèves doivent être capables se rapporte toujours à des éléments de contenu et à des processus ou des actions.

Le modèle de compétences HarmoS comprend quatre dimensions: la dimension du développement, les aspects de compétence, les domaines de compétence et les niveaux de compétence (niveaux d'exigence).

La dimension du développement décrit les compétences pour la fin de la 2^e année, de la 6^e année et de la 9^e année d'école.

Les huit aspects de compétence décrivent les aptitudes nécessaires pour résoudre des problèmes mathématiques:

- Savoir, reconnaître, décrire
- Appliquer des procédures, utiliser des techniques
- Utiliser des instruments et des outils
- Formuler, représenter

- Modéliser
- Argumenter, justifier
- Analyser, interpréter des résultats
- Explorer, essayer

Les cinq domaines de compétence décrivent les domaines de contenu ci-après des mathématiques (enseignées à l'école):

- Géométrie
- Nombres et opérations
- Fonctions
- Grandeurs et mesures
- Analyse de données

Le premier domaine recouvre les notions de formes et d'espace, le deuxième l'algèbre, le troisième les relations de dépendance, le quatrième le calcul pratique (résolution de problèmes) et le cinquième la stochastique (calcul des probabilités).

Les niveaux de compétence (I, II, III, IV) sont caractérisés par des tâches de différents degrés de difficulté, dont le niveau définitif est attribué en fonction de tests empiriques. Le degré de difficulté des tâches est déterminé en fonction des trois critères suivants:

- Difficulté liée à l'énoncé
- Difficulté liée à la résolution de la tâche (complexité)
- Difficulté liée à la familiarité de la tâche (plan d'études)

Les huit aspects de compétence appliqués aux cinq domaines de compétence aboutissent à une matrice de 40 champs, dans un premier temps vides, qu'il s'est agi de caractériser (voir figure III.15 pour la 9^e année). Chaque champ décrit ce que les élèves de l'année en question sont censés être capables de faire. Ces «*can do*» complétés par une énumération concrète des aptitudes et savoir-faire requis ont servi de base pour l'élaboration des exercices et des items des quatre niveaux de compétence.

Figure III.15: Aspects et domaines de compétence dans HarmoS (9^e année)

		Dimension du contenu («domaines de compétence»)				
		Nombres et opérations	Géométrie	Grandeurs et mesures	Fonctions	Analyse de données
Dimension des processus («aspects de compétence»)	Savoir, reconnaître, décrire					
	Appliquer des procédures, utiliser des techniques					
	Utiliser des instruments et des outils					
	Formuler, représenter					
	Modéliser					
	Argumenter, justifier					
	Analyser, interpréter des résultats					
	Explorer, essayer					

3.2.4 Le modèle PISA de compétences en mathématiques

Le modèle de compétences PISA est tridimensionnel (Klieme, Neubrand & Lüdtke, 2001/2002). Il distingue entre les contenus, les processus organisés par classes de compétences et les degrés d'exigence ou de difficulté.

Les idées majeures du contenu mathématique sont les suivantes:

- Quantité
- Variations et relations
- Espace et formes
- Incertitude

Les compétences sont les suivantes:

- Communication mathématique
- Modélisation
- Argumentation mathématique
- Utilisation d'un langage et d'opérations de nature symbolique, formelle et technique
- Représentation
- Création et résolution de problèmes

Ces domaines et compétences présentent des analogies avec le modèle du NCTM (voir section 3.2.1).

Chaque domaine de prestation, représenté par une cellule de la matrice des contenus et des compétences, est échelonné selon cinq niveaux ou degrés d'exigence/de difficulté:

- Degré I: Calcul au niveau de l'école primaire
- Degré II: Modélisation élémentaire
- Degré III: Modélisation, intégration et mise en relations au niveau secondaire I
- Degré IV: Modélisations avancées faisant appel à des notions complexes
- Degré V: Modélisation complexe et argumentation mathématique

Les descripteurs des niveaux d'exigence se réfèrent à quatre éléments de contenu (voir Klieme, Neubrand & Lüdtke, 2001, p. 160):

- Connaissances et contenus
- Exigences de modélisation
- Intégration et mise en relations
- Forme ouverte de la tâche

3.2.5 Modèles TIMSS

Différents modèles de compétences ont été développés dans le cadre des Etudes des tendances internationales en mathématiques et en sciences (TIMSS). L'enquête qui nous intéresse plus particulièrement ici est TIMSS/III, qui étudie les compétences en mathématiques et en physique en fin d'études secondaires et se réfère donc aux mathématiques et à la physique enseignées au gymnase (Ramseier et al., 1999, p. 130). Un aspect de l'étude est la tentative de caractériser les facteurs de difficulté des tâches de mathématiques et de physique.

L'enquête TIMSS/III s'est fondée sur un modèle de compétences plus différencié que ceux utilisés pour les populations d'élèves des classes inférieures. Ce modèle est décrit de façon détaillée par Baumert, Bos & Lehmann (2000).

Il comporte quatre niveaux d'aptitudes attendues:

- Reproduction de connaissances
- Maîtrise de procédures de routine
- Maîtrise de procédures complexes
- Tâches orientées vers les applications et problèmes de réflexion mathématique

Le modèle distingue également quatre niveaux d'exigence:

- Niveau 1: Conclusions portant sur la vie quotidienne
- Niveau 2: Application de routines simples
- Niveau 3: Modélisation et mise en relations d'opérations
- Niveau 4: Argumentation mathématique

Les contenus mathématiques sont structurés en cinq domaines:

- Nombres, équations et fonctions
- Analyse
- Géométrie, géométrie analytique
- Probabilités, statistique
- Logique des propositions, démonstrations

Les contenus et les aptitudes attendues peuvent être présentés sous la forme d'une matrice (figure III.16).

Figure III.16: Matrice des aptitudes attendues et des contenus

		Aptitudes attendues			
		Connaissances	Maîtrise de procédures de routine	Maîtrise de procédures complexes	Tâches orientées vers les applications et problèmes de réflexion mathématique
Contenus	Nombres, équations et fonctions				
	Analyse				
	Géométrie, géométrie analytique				
	Probabilités, statistique				
	Logique des propositions, démonstrations				

Un système de catégories supplémentaire a été développé spécialement pour les questions du test en mathématiques et en sciences dans TIMSS/III. Il distingue huit caractéristiques communes (variables). Ces caractéristiques constituent des exigences spécifiques en termes de connaissances, de compréhension, de processus réflexifs et d'aptitudes cognitives de base:

- G1. Connaissance de définitions, de propositions mathématiques (La tâche exige-elle de reproduire de mémoire de tels contenus de connaissance?)

- G2. Compréhension qualitative de notions mathématiques (La tâche exige-t-elle par exemple d'avoir compris ce qui relie un concept à d'autres concepts, dans quels cas il est applicable et non, et ce qui le différencie de notions employées dans le langage courant?)
- G3. Calcul (opérations arithmétiques)
- G4. Maniement de termes et de formules mathématiques
- G5. Interprétation de diagrammes («diagrammes» ici non pas au sens d'esquisses illustratives ou de constructions géométriques, mais au sens de représentations de coordonnées ou de diagrammes statistiques)
- G6. Compréhension de textes (La tâche exige-t-elle de lire et de comprendre un texte d'une certaine longueur ou des formulations linguistiques compliquées ou inhabituelles?)
- G7. Représentation visuelle
- G8. Processus de résolution de problèmes

Pour les exercices de mathématiques, quatre dimensions complémentaires sont proposées:

- M1. Formalisation (description de relations représentées sous forme de séries de données (mesures), de texte ou de graphes, sous forme algébrique, c'est-à-dire sous forme de termes, d'équations et de systèmes d'équation)
- M2. Interprétation du contexte (La tâche fait-elle référence à un contexte extra-mathématiques déterminant? Le répondant doit-il comprendre et interpréter ce contexte d'application? Doit-il pour cela disposer de connaissances extra-mathématiques?)
- M3. Restructuration/flexibilité (Cette caractéristique d'exigence est considérée comme significative lorsque la résolution de la tâche exige de reformuler/réorganiser d'une autre manière, non évidente au premier abord, une situation décrite verbalement ou une autre représentation.)
- M4. Principes de pensée mathématique (Il s'agit en l'occurrence quasiment de la pensée «métamathématique». Le répondant doit-il par exemple distinguer entre des concepts, des définitions, des appellations formelles et des interprétations (exemples)? Doit-il vérifier si une affirmation est généralisable ou valider un modèle mathématique? Doit-il avoir intégré l'obligation de prouver des affirmations mathématiques et en avoir compris la raison?)

Il s'est avéré que les questions de test développées pour EVAMAR II s'insèrent parfaitement dans ce modèle général, même si les catégories M1 à M4 se sont avérées trop abstraites et trop larges. Les domaines de contenu mathématique, en revanche, coïncident dans une large mesure avec les contenus identifiés pour l'enquête EVAMAR II (voir figure III.16).

3.3 Modèle final de la grille de compétences «mathématiques» pour EVAMAR II

A la différence d'autres projets comme HarmoS où les questions de test étaient élaborées exclusivement sur la base des modèles de compétences, la grille de compétences ne constitue qu'un élément de fond pour l'élaboration des exercices dans EVAMAR II. Dans la présente étude, les exercices doivent en effet se référer à des contenus (unités de sens) relevés de façon empirique dans des supports d'enseignement (notes de cours, manuels, etc.) d'universités suisses (voir section 1). La grille de compétences en mathématiques ne se fonde donc pas seulement sur une «modélisation théorique» développée à partir d'une analyse de modèles de compétences existants, mais s'est construite et développée parallèlement aux questions de test élaborées sur la base des unités de sens en mathématiques identifiées dans le module A. La dimension se rapportant au contenu mathématique a été ajustée aux unités de sens codées, et la dimension se rapportant aux compétences mathématiques n'a acquis sa forme définitive qu'après une première série de questions de test.

Les compétences telles que l'argumentation, la mathématisation, la modélisation ou la résolution de problèmes ne sont pas au centre de l'intérêt d'EVAMAR II, qui vise à évaluer dans quelle mesure les bacheliers et les bacheliers sont capables de comprendre les supports de cours des universités suisses sous l'angle des contenus mathématiques. De ce fait, la présente grille de compétences s'écarte des modèles de compétences décrits plus haut de la KMK et des projets HarMoS et PISA, lesquels donnent une place centrale à ce type de compétences. Par ailleurs, la grille de compétences pour EVAMAR II doit permettre de modéliser les exigences en termes de connaissances et de savoir-faire en mathématiques à l'interface gymnase/université. Il s'est avéré trop compliqué d'adapter des modèles de compétences conçus pour les niveaux inférieurs, ce qui disqualifiait les modèles de la KMK, de HarMoS et de PISA également sur le plan des contenus.

Parmi les modèles de compétences analysés, celui de TIMSS/III (Baumert et al., 2000; voir section 3.2.5) est le seul qui s'attache à décrire les compétences des élèves à la fin du degré secondaire II, soit au moment de passer la maturité. Les catégories G1 à G7 de ce modèle, qui s'appliquent aux questions du test, sont particulièrement intéressantes pour EVAMAR II. L'expérience a montré que les questions de test développées pour EVAMAR II s'insèrent parfaitement dans ce canevas. Par opposition, les catégories M1 à M4 se sont avérées trop abstraites et trop larges. Les domaines de contenu de TIMSS/III, en revanche, se recouvrent à nouveau très bien avec les contenus répertoriés dans EVAMAR II (voir figure III.16). Le modèle TIMSS/III a donc servi de base pour la présente grille de compétences.

Pour la dimension des contenus, il est apparu approprié de ne distinguer que quatre domaines thématiques au lieu des cinq du modèle TIMSS, soit «Nombres, équations et fonctions», «Analyse», «Géométrie et géométrie analytique», «Probabilités et statistique» et «Logique des propositions et démonstrations». Le module A n'a répertorié qu'un nombre insuffisant d'unités de sens se rapportant au domaine «Logique des propositions et démonstrations». Par ailleurs, l'analyse des supports d'enseignement a mis au jour des unités de sens liées à la théorie des ensembles et à l'algèbre élémentaire. Il a donc été décidé de définir pour EVAMAR II, au titre de la dimension principale du contenu, les quatre domaines principaux «Calcul analytique», «Géométrie», «Stochastique» et «Divers», ce dernier domaine regroupant tout ce qui ne se rattache pas aux trois autres. La distinction entre ces quatre domaines principaux se prête aussi bien aux contenus mathématiques enseignés au gymnase qu'à ceux traités à l'université. Cet aspect a joué un rôle dans la perspective du sous-projet D1 (Analyse d'examens de maturité). De plus, la limitation à quatre domaines permet de prévoir plus de tâches par catégories (les «cellules» de la matrice), ce qui est avantageux pour l'interprétation statistique du test. Après l'analyse des unités de sens identifiées dans le sous-projet A1 et l'élaboration des questions du test, chacun des quatre domaines principaux a été lui-même divisé en trois sous-domaines précisant son contenu.

Au final, la dimension du contenu de la grille de compétences développée pour EVAMAR II présente la structure suivante:

- Calcul analytique:
 - Fonctions
 - Calcul infinitésimal (calcul différentiel et intégral)
 - Suites et séries
- Géométrie:
 - Géométrie élémentaire
 - Trigonométrie
 - Géométrie vectorielle
- Stochastique:
 - Calcul des probabilités
 - Analyse combinatoire
 - Statistique

- Divers:
 - Théorie des ensembles, mathématiques discrètes
 - Logique
 - Algèbre élémentaire

La dimension des compétences de la grille a également été définie après seulement qu'une première série de questions de test ont été élaborées. Là encore, l'expérience a montré que le système de catégories développé pour TIMSS/III se prêtait très bien à cette fin. Il s'agit des compétences suivantes (cf. Baumert et al., 2000):

- Connaissance de définitions et de propositions mathématiques: il suffit de connaître ou de reconnaître, par exemple, la formule du binôme ou le théorème de Pythagore dans les supports de cours.
- Compréhension qualitative de notions mathématiques: cette catégorie comprend des tâches pour lesquelles la simple connaissance d'une définition ou d'une loi mathématique n'est pas suffisante.
- Calcul: cette catégorie comprend des tâches qui se limitent au maniement de nombres. Des variables, des inconnues ou des termes avec variables n'y apparaissent pas.
- Maniement de termes et de formules mathématiques: appartiennent à cette catégorie des tâches faisant intervenir des termes en partie autres que des nombres.
- Interprétation de diagrammes: il s'agit en l'occurrence de l'interprétation de représentations de données statistiques et de représentations mathématiques comme des courbes de fonctions ou des dessins géométriques.
- Compréhension de textes: certaines tâches sont particulièrement exigeantes en termes de compréhension textuelle. Les tâches rangées dans cette catégorie appartiennent souvent en même temps à une autre catégorie.
- Représentation visuelle: la résolution de tâches de cette catégorie requiert une capacité de visualisation ou d'orientation dans l'espace à trois dimensions.

La grille de compétences élaborée pour EVAMAR II est présentée dans la figure III.17.

Figure III.17: Grille de compétences en mathématique pour EVAMAR II

	CALCUL ANALYTIQUE			GEOMETRIE			STOCHASTIQUE			DIVERS		
	Fonctions	Calcul infinitésimal	Suites et séries	Trigonométrie	Géométrie analytique	Calcul vectoriel	Calcul des probabilités	Analyse combinatoire	Statistique	Théorie des ensembles, mathématiques discrètes	Logique	Algèbre élémentaire
Connaissance de définitions et de propositions												
Compréhension qualitative de notions mathématiques												
Calcul												
Maniement de termes et de formules mathématiques												
Interprétation de diagrammes												
Compréhension de textes												
Représentation visuelle												

4 La grille de compétences en biologie

4.1 Bases

La grille de compétences en biologie se fonde elle aussi sur une analyse de descripteurs de compétences existants pour cette discipline. Dans la suite du projet EVAMAR II, cet outil servira également pour l'élaboration, sur le plan du contenu, des items du test de biologie (module C) et pour la catégorisation des examens de maturité en biologie (sous-projet D1) sur les plans du contenu et, éventuellement, des niveaux d'exigences cognitives. Dans la perspective de l'utilisation de cette grille, il fallait tenir compte de la limitation due au fait que le module C ne permettrait pas, en raison de contraintes méthodologiques et de ressources insuffisantes, de tester les performances dans certaines compétences opérationnelles, par exemple en matière d'acquisition des connaissances, de communication et d'évaluation. Etant donné qu'un modèle de compétences devrait idéalement considérer l'ensemble des compétences opérationnelles (Klieme, 2003; Weinert, 2001), le module C ne peut prétendre réaliser un véritable test de compétences, mais plutôt un test de connaissances. En conséquence, il est plus approprié de désigner le présent modèle comme «grille de connaissances spécifiques en biologie».

La section 4.2 ci-dessous passe en revue quelques classifications de la biologie en sous-domaines existant aux niveaux suisse et international. Suit un aperçu des dimensions proposées pour spécifier les niveaux d'exigences cognitives requises pour la résolution de problèmes de biologie. Ces réflexions aboutissent enfin dans une grille de connaissances bidimensionnelle dans la discipline de la biologie (section 4.3).

4.2 Grilles et modèles de compétences existants pour la biologie

En Suisse, on trouve des éléments de description des connaissances spécifiques exigées par disciplines au degré secondaire II dans les directives 2003-2006 pour l'examen suisse de maturité (validité prolongée jusqu'en 2008), élaborées au niveau fédéral (CSM, 2008). Ces directives constituent les seuls éléments de compétences définis spécialement pour la Suisse dans le domaine des sciences expérimentales au degré secondaire II. Elles décrivent notamment les objectifs de formation et le programme de la discipline fondamentale Biologie. Elles sont comparables aux objectifs d'apprentissage que chaque corps enseignant en Suisse, typiquement, définit pour sa propre école et applique aux examens de maturité organisés dans son établissement. Cependant, le projet EVAMAR II ayant également pour objet d'évaluer les examens de maturité en soi (sous-projet D1), ces directives doivent être écartées d'emblée de la modélisation de la grille de compétences en biologie et ne sont donc plus mentionnées dans l'aperçu ci-dessous.

Il existe par ailleurs des systèmes de classification spécialement axés sur les contenus scientifiques et constamment mis à jour, dont le plus connu dans l'espace germanophone est certainement la classification allemande de Ratisbonne, ou RVK (*Regensburger Verbundklassifikation*).

4.2.1 La classification allemande de Ratisbonne (RVK)

On constate rapidement que de telles classifications répondent en premier lieu à des besoins de classement systématique des matières dans les bibliothèques et non à des critères didactiques ou pédagogiques. Pour les bibliothèques, des catégories telles que «Revue et résumés» ou «Bibliographies et ouvrages de référence», «Histoire et didactique» ou «Manuels de biologie» font sens, alors qu'elles ne font naturellement pas partie des matières scolaires.

Hormis ces domaines particuliers, la classification de Ratisbonne distingue les sous-domaines de la biologie suivants:

Cytologie; microbiologie; biotechnologie; immunologie; génétique; évolution; écologie; botanique spéciale (systématique); morphologie et anatomie des plantes; physiologie des plantes; zoologie spéciale (taxinomique); recherche comportementale et psychologie animale; anthropologie; morphologie, anatomie et physiologie

générales, anatomie, physiologie et biochimie de certains organes et systèmes organiques humains et animaux; alimentation, digestion, métabolisme, équilibre énergétique, reproduction, physiologie du développement humain et animal, physiologie appliquée.

Comme on le verra plus bas, il est relativement aisé de relier ces catégories aux classifications du domaine pédagogique scolaire (voir figure III.21, dernière ligne). Une exception à cet égard est le sous-domaine «Anthropologie» avec sa subdivision dans une branche plutôt matérialiste-physique («Quelle phylogenèse a conduit à l'être humain?») et une branche plutôt idéaliste-pragmatique («Qu'est-ce qui fait de l'être humain un être humain?»). Tandis que la branche matérialiste-physique peut clairement être rangée dans le domaine de l'évolution, la branche idéaliste-pragmatique nécessiterait de fait la définition d'un sous-domaine à part.

Un premier enseignement tiré de l'observation de la RVK est que la grille de connaissances spécifiques en biologie développée pour la présente étude s'avère très large et à même d'englober dans son amplitude toutes les facettes de la biologie à l'exception des contenus relevant de la branche idéaliste-pragmatique de l'anthropologie. Dans le contexte de l'«aptitude générale aux études» qui est au cœur de la présente évaluation, on peut toutefois supposer que ces contenus auraient eu plus d'importance si l'on avait étudié la discipline de la philosophie. Les résultats du module A n'apportent aucun élément indiquant que des contenus relevant de la tendance idéaliste-pragmatique de l'anthropologie auraient eu une quelconque importance pour l'évaluation de l'aptitude aux études dans les disciplines prises en compte, en particulier pour les études de biologie.

Le reste du présent chapitre analyse dans quelle mesure les classifications issues de la pédagogie scolaire se prêtent aux objectifs de la présente évaluation. Pour la définition des contenus, on se fondera principalement sur:

1. l'analyse de contenu de manuels de biologie standard de langue allemande pour le niveau gymnasial (secondaire II) (section 4.2.2) et celle de deux ouvrages standard étasuniens pour les étudiants débutant à l'université (section 4.2.3),
2. les exigences communes pour l'examen de baccalauréat allemand en biologie (EPA Biologie) et les standards de formation en Allemagne (section 4.2.4).

Au niveau des niveaux d'exigences cognitives, on s'intéressera en particulier:

1. à l'introduction d'une dimension spécifique de la «culture scientifique» («Scientific Literacy») proposée dans les pays anglosaxons (section 4.2.6),
2. aux exigences communes pour l'examen de baccalauréat allemand en biologie (EPA Biologie) et aux standards de formation en Allemagne (section 4.2.7 et 4.2.8).

4.2.2 Comparaison de la structure de manuels standard pour l'enseignement gymnasial en allemand

En raison de leur orientation vers des objectifs didactiques et pédagogiques spécifiques et de leur élaboration souvent étalée sur de longues années avec de nombreuses révisions, on peut attendre d'un manuel standard une structure du champ de connaissances particulièrement cohérente et raisonnée et particulièrement significative pour le domaine d'application de la pédagogie scolaire concerné dans le cadre d'EVAMAR II (voir les critères à la section 1). La structure des manuels en chapitres est un précieux point de départ pour structurer la biologie en domaines de connaissances.

La figure III.18 compare la structure en chapitres de trois manuels scolaires spécifiquement destinés à l'enseignement de la biologie au degré secondaire II et se prêtant donc particulièrement bien pour l'enseignement à la fin de la formation gymnasiale. Ces ouvrages sont utilisés de façon régulière en Suisse depuis plusieurs décennies. Le plus connu d'entre eux est probablement le manuel «Linder Biologie» (2005). On constate d'emblée un haut degré de convergence entre ces ouvrages, qui peut bien entendu être interprétée comme l'expression d'une convention peu réfléchie. Il est cependant beaucoup plus vraisemblable que cette similarité soit l'aboutissement d'une convergence qui s'est imposée naturellement au cours de l'évolution des structures imaginables et proposées au fil du temps. Les différences les plus marquées apparaissent lorsque les ouvrages consacrent des chapitres à part à des domaines de théorie scientifique particuliers ou à des questions

spécifiques liées aux applications. La plus grande analogie apparaît au contraire dans les chapitres consacrés à la cytologie, au métabolisme, à l'immunobiologie, à l'écologie et à la biologie de l'évolution. Les sous-domaines de la génétique, de la reproduction et de la biologie du développement, qui se recouvrent partiellement, sont groupés de différentes manières. A l'exception du volume complet «Biologie Oberstufe» (Cornelsen Verlag), tous les ouvrages étudiés consacrent un chapitre propre à la biologie du développement. Il en va de même pour la biologie comportementale, également à l'exception de «Biologie Oberstufe». Les avis varient sur la question de savoir s'il est plus approprié pour un manuel scolaire de souligner les relations transversales à l'intérieur de grands chapitres, comme le fait l'édition Cornelsen, ou de privilégier la clarté de la vue d'ensemble en divisant la matière en plusieurs petits chapitres, à l'instar des autres éditeurs. Dans le contexte d'EVAMAR II, la finalité principale de la grille de compétences qu'il s'agit de développer consiste à attribuer avec certitude un champ thématique spécifique aux questions de l'enquête principale et à un échantillon d'examens de maturité. Sachant que les problèmes de biologie, en particulier les problèmes au potentiel d'activation cognitive élevé, font souvent simultanément référence à plusieurs domaines thématiquement voisins, le travail de codage sera simplifié si les domaines thématiquement proches sont d'emblée regroupés dans des catégories relativement larges.

4.2.3 Comparaison de la structure de manuels standards pour les étudiants débutant en biologie

Les deux ouvrages de synthèse pour les études de base en biologie (Campbell & Reece, 2006 et Purves et al., 2008) présentent une concordance relativement élevée entre eux, mais une faible correspondance avec la structure en chapitres des manuels du degré secondaire II (figure III.19). Fait remarquable, les deux manuels traitent ensemble dans un même chapitre le domaine thématique «Métabolisme et échanges énergétiques» et celui de la «Communication intracellulaire», du «Traitement neuronal des informations» et de la «Perception». Les chapitres thématiquement très larges qui résultent de cette manière de voir englobent ainsi tous les aspects liés à la compréhension de la fonction d'un organisme (ou d'un organe) considéré globalement. La division en chapitres ne se fait pas selon des notions de base de la biologie, mais selon une division propre à la systématique, c'est-à-dire la distinction entre les plantes et les animaux. Les domaines restants, qui appartiennent soit au niveau clairement infra-organismique (chimie de la vie, cellules isolées, génétique) soit au niveau clairement supra-organismique (écologie ou théorie de l'évolution) sont traités dans des chapitres séparés.

Figure III.18: Comparaison de la structure de trois manuels standard de biologie en langue allemande au niveau gymnasial (degré secondaire II)
(titres originaux des chapitres en allemand)

	Erkenntnis wege der Biologie	Cytologie	Stoff- wechsel und Energie- haushalt	Genetik	Entwicklungsbiologie	Immunbiologie	Ökologie	Neuro- biologie	Hormone	Verhaltens- biologie	Evolution	Baupläne der Lebewesen	--
Linder Biologie (2005): Gesamtband SII, 22. Auflage. Schroedel, Braunschweig.													
Biologie heute entdecken (2004). Allgemeine Ausgabe für SII. Schroedel, Braunschweig.	--	Zell- biologie	Stoffwechsel	Genetik	Fortpflanzung und Entwicklung	Immunbiologie	Ökologie	Nerven-, Sinnes- und Hormonphysiologie		Verhaltens- biologie	Evolution- biologie	Das System der Lebewesen	Angewandte Biologie
Biologie - Oberstufe (2001). Gesamtband. Cornelsen, Berlin.	Biologie - die Wissen- schaft vom Leben	Zell- biologie	Stoffwechsel	Vererbung, Fortpflanzung und Entwicklung	Fortpflanzung und Entwicklung	Immunbiologie	Ökologie	Informationsverarbeitung, Regelung und Verhalten		Evolution	--	--	--

Figure III.19: Comparaison de la structure de deux manuels standard américains et traduits en allemand destinés aux étudiants débutant en biologie (degré tertiaire)
(titres originaux des chapitres en anglais et en allemand)

Campbell, N.A. & Reece, J.B. (2005): Biology, 7 th edition. Pearson Education; Upper Saddle River.	The Chemistry of Life	The Cell	Genetics	--	Mechanisms of Evolution	The Evolutionary History of Biological Diversity	Plant Form and Function	Animal Form and Function	Ecology
Campbell, N.A. & Reece, J.B. (2006): Biology, 6. Auflage Pearson Education, München.	Die Chemie des Lebens	Die Zelle	Die Gene	--	Die Mechanismen der Evolution	Die Stammesgeschichte der biologischen Diversität	Form und Funktion der Pflanzen	Form und Funktion der Tiere	Ökologie und Verhalten
Purves, W.K. et al. (2008): Life - The Science of Biology, 8 th edition. Sinauer Associates, Sunderland.	The Science and Building Blocks of Life	Cells and Energy	Heredity and the Genome The Genome in Action	--	The Patterns and Processes of Evolution	The Evolution of Diversity	Flowering Plants: Form and Function	Animals: Form and Function	Ecology
Purves, W.K. et al. (2006): Biologie, 7. Auflage. Spektrum Akademischer Verlag, München.	--	Die Zelle	Information und Vererbung	Entwicklung	Evolutionsprozesse	Die Evolution der biologischen Vielfalt	Die Biologie der Blütenpflanzen	Die Physiologie der Tiere	Ökologie und Biogeographie

4.2.4 Les EPA Biologie et les standards de formation en Allemagne

Les textes de référence traités dans la présente section sont désignés ci-après par leur dénomination courante abrégée selon le glossaire de la figure III.20.

Figure III.20: Nom court et contexte des documents traités ci-après

Nom court	Contexte
Standards de formation KMK en biologie	Standards de formation pour la fin de l'école moyenne (secondaire I) en biologie, définis par décision du 16.12.2004 de la Conférence permanente des ministres de l'éducation et des affaires culturelles des länder allemands (KMK pour <i>Kultusministerkonferenz</i>) (KMK, 2005)
EPA Biologie	Exigences communes pour l'examen de baccalauréat (secondaire II) en biologie, définies par décision du 1.12.1989 de la KMK, état le 5.2.2004 (EPA Biologie pour <i>Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie</i>) (KMK, 2004)
Avis de la MNU	Avis, rendu sous forme de lettre, de la Société allemande pour la promotion de l'enseignement des mathématiques et des sciences naturelles (MNU, <i>Deutscher Verein zur Förderung des mathematischen und naturwissenschaftlichen Unterrichts e.V.</i>) sur les projets de standards de la formation pour la fin de l'école secondaire en biologie, chimie et physique, état le 30.8.2004

Des chercheurs du Leibniz-Institut für die Pädagogik der Naturwissenschaften à Kiel (IPN) développent actuellement dans le cadre du projet «Biologie en contexte» (bik), réalisé sur mandat du Ministère fédéral allemand de la formation et de la recherche (BMBF), un modèle de compétences en biologie pour le degré secondaire I (SI). Il est prévu de fixer les niveaux de compétences de la dimension des connaissances spécifiques de façon empirique au moyen de questions-test. A l'heure de la rédaction du présent rapport (août 2008), les questions-test sont en phase de construction et il est prévu de les évaluer prochainement sur une population du degré SI (communication personnelle de Philippe Schmiemann, Universität Duisburg). Cette définition en cours de la dimension des «connaissances spécifiques» s'appuiera très vraisemblablement sur les précédents travaux des standards de formation KMK en biologie (2005), eux aussi relatifs au degré secondaire I (cf. section 4.2.4.2).

4.2.4.1 EPA Biologie

Les EPA Biologie distinguent pour la fin du degré secondaire II deux «perspectives» pertinentes pour la présente étude: il s'agit d'une part des «Domaines thématiques» et d'autre part des «Notions de base» (la troisième perspective, «Eléments de réflexions sur la conception de l'homme», ne nous intéresse pas ici, cf. section 4.2.1). Ces «perspectives» constituant différents points de vue d'un même objet, il est logique que leurs contenus soient relativement étroitement imbriqués l'un dans l'autre.

Perspective thématique «Domaines thématiques»

La première perspective pertinente pour une division de la biologie en sous-domaines est celle des domaines thématiques (KMK, 2004, chap. 1.2.1. Themenbereiche, p. 7 s):

Relations fonctionnelles et bases moléculaires de ces dernières

Thèmes liés à la physiologie, à la biologie cellulaire, à la génétique

Des relations visibles à l'échelle macroscopique aux relations microscopiques et moléculaires: une analyse toujours plus approfondie des relations de cause à effet stimule les processus cognitifs à différents niveaux et ouvre la compréhension aux phénomènes complexes, par déduction et généralisation.

Font partie du domaine thématique Relations fonctionnelles:

- Structure et fonction des cellules, tissus et organes; différenciations fonctionnelles
- Communication intercellulaire
- Traitement neuronal de l'information, perception
- Bases de la génétique moléculaire
- Métabolisme et échanges énergétiques: nécessité et formes de transformations de l'énergie, assimilation et dissimilation dans les relations entre les structures cellulaires et l'organisme
- Catalyseurs moléculaires du métabolisme: enzymologie
- Applications de connaissances et de méthodes biologiques modernes, par ex. technologie génique, biologie de la reproduction, biotechnologie

Systèmes interdépendants - écologie et développement durable

L'écologie thématise les interactions entre l'organisme et son environnement ainsi que le tissu de relations entre les niveaux d'organisation de la population et de l'écosystème. Le rapport de l'homme à la nature en fait également partie.

Font partie du domaine thématique Ecologie et développement durable:

- Etude et analyse d'un écosystème, par ex. études d'inventaire, procédés de mesure
- Facteurs écologiques, biotope et biocénose
- Biodiversité et développement des populations
- Cycle des éléments et flux d'énergie
- Êtres humains et écosystèmes: l'influence de l'homme sur l'écosystème et ses effets, conséquence pour l'homme et pour l'humanité

Processus de développement - évolution et questions d'avenir

L'évolution reflète la diversité des êtres vivants et de leurs interactions.

La perspective de l'évolution éclaire l'origine et l'évolution du vivant en tant que processus phylogénétique, dont nous faisons partie en tant qu'êtres humains. Elle inclut aussi des questions sur l'évolution future de la vie sur Terre.

Font partie du domaine thématique Evolution et questions d'avenir:

- Phylogénèse, liens de parenté, systématique
- Origine des formes et des espèces, variabilité et construction de niche
- Adaptation, y compris questions de sociobiologie
- Théories de l'évolution, appréciation de leur signification
- Origine et avenir de l'espèce humaine

Perspective thématique «Notions de base»

Les EPA Biologie distinguent les notions de base ci-après (KMK, 2004, chap. 1.2.2. Basiskonzepte, p. 8):

a) Structure et fonction

Les organismes vivants et les processus de la vie répondent à des structures; structure et fonction sont liées.

Cette notion de base aide par exemple à comprendre la structure des biomolécules, la fonction des enzymes, des organes et des écosystèmes.

b) Reproduction

Les organismes vivants sont capables de se reproduire; la reproduction implique la transmission d'informations héréditaires.

Cette notion de base aide par exemple à comprendre la réplication identique de l'ADN, les virus, la mitose et la reproduction sexuelle.

c) Compartimentation

Les systèmes vivants fonctionnent et réagissent dans un espace limité.

Cette notion de base aide par exemple à comprendre les organites des cellules, les organes et la biosphère.

d) Contrôle et régulation

Les systèmes vivants maintiennent certains états par des mécanismes de régulation et réagissent aux changements auxquels ils sont soumis.

Cette notion de base aide par exemple à comprendre la biosynthèse des protéines, la régulation hormonale et l'évolution des populations.

e) Métabolisme et transformation d'énergie

Les organismes vivants sont des systèmes ouverts, qui impliquent des processus métaboliques et de transformation d'énergie.

Cette notion de base aide par exemple à comprendre la photosynthèse, la nutrition et le transport de substances et leurs transformations successives dans l'organisme.

f) Information et communication

Les organismes vivants reçoivent des informations, les enregistrent, les traitent et communiquent.

Cette notion de base est par exemple utile pour comprendre le codage d'informations au niveau des macromolécules, la transmission de stimuli, l'apprentissage et le comportement territorial.

g) Variabilité et adaptation

Les organismes vivants sont adaptés à leur environnement dans leur structure et leur fonction. L'adaptation est possible grâce à la variabilité. Les bases de la variabilité chez les organismes vivants sont la mutation, la recombinaison et la modification.

Cette notion de base est par exemple utile pour comprendre l'anémie à cellules falciformes, les niches écologiques et la spéciation.

h) Histoire et parenté

Les analogies et la diversité des organismes vivants sont le résultat de processus d'évolution phylogénétique.

Cette notion de base aide par exemple à comprendre l'apparition du vivant, les organes homologues et l'origine de l'homme.

Etant donné que ces «Domaines thématiques» et ces «Notions de base» constituent différentes perspectives d'un seul et même objet, on aurait pu envisager de les représenter dans une matrice en faisant systématiquement correspondre, en leur point de croisement, les deux perspectives portant sur un même objet. Un tel degré de compartimentation thématique n'est cependant pas utile en l'occurrence. Nous avons donc préféré rattacher les notions de base aux domaines thématiques auxquels elles sont le plus souvent associées (voir figure III.21).

4.2.4.2 Standards de formation KMK en biologie

Sans référence explicite aux EPA Biologie ni à toute autre source et malheureusement sans explications détaillées de leur approche, les auteurs des **standards de formation KMK en biologie** proposent trois notions de base pour la fin du degré secondaire I (KMK, 2005, chap. 2.1, p. 8 ss):

- a) Système
- b) Structure et fonction
- c) Développement

Cette réduction à trois notions de base n'apporte pas de valeur ajoutée pour EVAMAR II par rapport aux EPA Biologie, et ce pour les deux raisons suivantes:

1. Argument du degré considéré: le grand avantage des **EPA Biologie** est qu'elles se réfèrent expressément au degré secondaire II, tandis que les **standards de formation KMK en biologie** se rapportent au secondaire I.
2. Argument du «recul qualitatif»: les auteurs partagent en l'occurrence **l'avis de la MNU** concernant la réduction du nombre de notions de base:

«La réduction de la dimension des contenus aux trois seules notions de base Système, Structure et fonction et Développement est trop générale et de ce fait non adaptée à la discipline de la biologie. La notion de «Système» (p. 8), en particulier, n'est dans son abstraction pas un concept adéquat pour permettre à des élèves du degré secondaire I d'analyser un contexte, de structurer des contenus et de les systématiser...» (p. 7 s)²¹.

«Les standards de formation pour la fin de la formation secondaire ne sont pas compatibles avec les EPA harmonisées (avec succès) au niveau fédéral (printemps 2004):

Des huit notions de base des EPA, une seule a été reprise à laquelle on a ajouté deux autres notions de niveau sensiblement plus abstrait.

Des contenus essentiels de la biologie sont laissés de côté, comme l'évolution/la sélection, la biologie humaine et la régulation.

Résultat, la biologie n'est pas représentée dans son entièreté, la base de la formation générale est réduite, **la biologie est limitée au champ de connaissances de la reproduction**, au détriment d'une compréhension nécessaire de la biologie atteignable grâce à un apprentissage cumulatif et tissé à partir de notions de base plus nombreuses et plus concrètes.

Sans le fondement théorique des notions d'évolution et de sélection, il n'est point de biologie ni d'enseignement sensé de la biologie. Ce déficit apparaît de façon très nette dans les explications sur la notion de Structure et fonction: le texte est imprécis, ne respecte en aucune façon les modes de raisonnement et d'explication propres à la biologie et néglige le fait qu'une explication fonctionnelle n'a de sens que si elle se fonde sur la théorie de l'évolution et de la sélection.

Les notions et champs conceptuels plus différenciés Reproduction, Variabilité et adaptation, Structure et fonction, Métabolisme et transformation d'énergie, Temps, Niveaux, Régulation, Interaction, Information, Spécificité de l'homme sont d'une part plus adaptés à la diversité thématique de la biologie, mais sont aussi plus concrets et plus évocateurs pour le public auquel ils s'adressent, soit les enseignants de biologie et les élèves du degré secondaire I (...).» (Mises en évidence dans l'original)²²

²¹ Citation originale en allemand, traduite en français pour la présente version du rapport.

²² Idem

4.2.5 Résultat intermédiaire pour la dimension des contenus: synthèse

Figure III.21: Classification de la biologie en sous-domaines, fondée sur des ouvrages de référence standard de biologie en langues allemande et anglaise, sur les EPA Biologie allemandes et sur la classification allemande de Ratisbonne

Source	Degré	Cytologie/ Anatomie/Métabolisme	Traitement de l'information/ Comportement	Immunobiologie	Génétique/ Biologie du développement	Ecologie	Systématique/- Evolution
Linder Biologie (2005): Gesamtband SII, 22. Ausgabe. Schroedel, Braunschweig.	SII	<ul style="list-style-type: none"> • Cytologie • Métabolisme et équilibre énergétique 	<ul style="list-style-type: none"> • Neurobiologie • Hormones • Comportement 	<ul style="list-style-type: none"> • Immunobiologie 	<ul style="list-style-type: none"> • Génétique • Biologie du développement 	<ul style="list-style-type: none"> • Ecologie 	<ul style="list-style-type: none"> • Structures des êtres vivants • Evolution
Biologie Oberstufe (2001). Gesamtband. Cornelsen, Berlin.	SII	<ul style="list-style-type: none"> • Biologie cellulaire • Métabolisme 	<ul style="list-style-type: none"> • Traitement de l'information, régulation et comportement 	<ul style="list-style-type: none"> • Immunobiologie 	<ul style="list-style-type: none"> • Héritéité, reproduction et développement 	<ul style="list-style-type: none"> • Ecologie 	<ul style="list-style-type: none"> • Evolution
Biologie heute entdecken (2004). Allgemeine Ausgabe für SII. Schroedel, Braunschweig.	SII	<ul style="list-style-type: none"> • Biologie cellulaire • Métabolisme 	<ul style="list-style-type: none"> • Nerven-, Sinnes- und Hormonphysiologie • Verhaltensbiologie 	<ul style="list-style-type: none"> • Immunobiologie 	<ul style="list-style-type: none"> • Génétique • Reproduction et développement • Biologie appliquée 	<ul style="list-style-type: none"> • Ecologie 	<ul style="list-style-type: none"> • Le système des organismes vivants • Biologie de l'évolution
Campbell, N.A. & Reece, J.B. (2005): Biology, 7 th edition. Pearson Education; Upper Saddle River.	tertiaire	<ul style="list-style-type: none"> • The Chemistry of Life (La chimie du vivant) • The Cell (La cellule) • Plant Form and Function (Forme et fonction des plantes) • Animal Form and Function (Forme et fonction des animaux) 			<ul style="list-style-type: none"> • Genetics (Les gènes) 	<ul style="list-style-type: none"> • Ecology (Ecologie et comportement) 	<ul style="list-style-type: none"> • Mechanisms of Evolution (Les mécanismes de l'évolution) • The Evolutionary History of Biological Diversity (La phylogénèse de la diversité biologique)
Campbell, N.A. & Reece, J.B. (2006): Biology, 6. Auflage. Pearson Education, München.							
Purves, W.K. et al. (2008): Life - The Science of Biology,	tertiaire	<ul style="list-style-type: none"> • The Science and Building Blocks of Life (--) • Cells and Energy (La cellule) • Flowering Plants: Form and Function (La biologie des plantes à fleurs) 			<ul style="list-style-type: none"> • Heredity and the Genome (Information et héritéité) 	<ul style="list-style-type: none"> • Ecology (Ecologie et biogéographie) 	<ul style="list-style-type: none"> • The Patterns and Processes of Evolution (The Evolution of

Source	Degré	Cytologie/ Anatomie/Métabolisme	Traitement de l'information/ Comportement	Immunobiologie	Génétique/ Biologie du développement	Ecologie	Systématique/- Evolution	
8 th edition. Sinauer Associates, Sunderland. Purves, W.K. et al. (2006): Biologie, 7. Auflage. Spektrum Akademischer Verlag, München		<ul style="list-style-type: none">Animals: Form and Function (La physiologie des animaux)			<ul style="list-style-type: none">The Genome in Action (Développement)		Diversity) <ul style="list-style-type: none">Evolutionprozesse (L'évolution de la diversité biologique)	
KMK (éd.) (2004): Sekretariat de la Conférence permanente des ministres de l'éducation et des affaires culturelles des länder de la République fédérale d'Allemagne (KMK): Exigences communes pour l'examen de baccalauréat en biologie (décision du 1.12.1989 de la KMK, état le 5.2.2004)	SII	A. Relations fonctionnelles et bases moléculaires de ces dernières Thèmes liés à la physiologie, à la biologie cellulaire, à la génétique Des relations visibles à l'échelle macroscopique aux relations microscopiques et moléculaires: une analyse toujours plus approfondie des relations de cause à effet stimule les processus cognitifs à différents niveaux et ouvre la compréhension aux phénomènes complexes, par déduction et généralisation. Font partie du domaine thématique Relations fonctionnelles: <ul style="list-style-type: none">Structure et fonction des cellules, tissus et organes; différenciations fonctionnellesCommunication intercellulaireTraitement neuronal de l'information, perceptionBases de la génétique moléculaireMétabolisme et échanges énergétiques: nécessité et formes de transformations de l'énergie, assimilation et dissimilation dans les relations entre les structures cellulaires et l'organismeCatalyseurs moléculaires du métabolisme: enzymologieApplications de connaissances et de méthodes biologiques modernes, par ex. technologie génique, biologie de la reproduction, biotechnologie <ul style="list-style-type: none">Remarque: l'immunobiologie n'est pas mentionnée expressément					B. Systèmes interdépendants – écologie et développement durable L'écologie thématise les interactions entre l'organisme et son environnement ainsi que le tissu de relations entre les niveaux d'organisation de la population et de l'écosystème. Le rapport de l'homme à la nature en fait également partie. Font partie du domaine thématique Ecologie et développement durable: <ul style="list-style-type: none">Etude et analyse d'un écosystème, par ex. études d'inventaire, procédés de mesureFacteurs écologiques, biotope et biocénoseBiodiversité et développement des populations	C. Processus de développement - évolution et questions d'avenir L'évolution reflète la diversité des êtres vivants et de leurs interactions. La perspective de l'évolution éclaire l'origine et l'évolution du vivant en tant que processus phylogénétique, dont nous faisons partie en tant qu'êtres humains. Elle inclut aussi des questions sur l'évolution future de la vie sur Terre. Font partie du domaine thématique Evolution et questions d'avenir: <ul style="list-style-type: none">Phylogénèse, liens de parenté, systématiqueOrigine des formes et des espèces, variabilité et

Source	Degré	Cytologie/ Anatomie/Métabolisme	Traitement de l'information/ Comportement	Immunobiologie	Génétique/ Biologie du développement	Ecologie	Systématique/- Evolution
						<ul style="list-style-type: none"> Cycles des éléments et flux d'énergie Etres humains et écosystèmes: l'influence de l'homme sur l'écosystème et ses effets, conséquence pour l'homme et pour l'humanité 	<ul style="list-style-type: none"> construction de niche Adaptation, y compris questions de sociobiologie Théories de l'évolution, appréciation de leur signification Origine et avenir de l'espèce humaine
		Sous-domaine 1 Structure et fonction des cellules, tissus et organes; différenciations fonctionnelles Notion de base connexe Compartmentation Les systèmes vivants fonctionnent et réagissent dans un espace limité. Notion de base connexe Structure et fonction Les organismes vivants et les processus de la vie répondent à des structures; structure et fonction sont liées. Sous-domaine 4 Métabolisme et échanges énergétiques: nécessité et formes de transformations de l'énergie, assimilation et	Sous-domaine 2 Communication intercellulaire, traitement neuronal de l'information, perception Notion de base connexe Information et communication Les organismes vivants reçoivent des informations, les enregistrent, les traitent et les communiquent. Notion de base connexe Contrôle et régulation Les systèmes vivants maintiennent certains états par des mécanismes de régulation et réagissent aux changements auxquels ils sont soumis. Notion de base connexe Reproduction Les organismes vivants sont capables de se reproduire; la reproduction implique la transmission d'informations héréditaires.	Non mentionnée expressément (!)	Sous-domaine 3 Bases de la génétique moléculaire, applications de connaissances et de méthodes biologiques modernes, par ex. technologie génique, biologie de la reproduction, biotechnologie Notion de base connexe Reproduction Les organismes vivants sont capables de se reproduire; la reproduction implique la transmission d'informations héréditaires.	Sous-domaine 5 Systèmes interdépendants – écologie et développement durable Notion de base connexe Métabolisme et transformation d'énergie Les organismes vivants sont des systèmes ouverts, qui impliquent des processus métaboliques et de transformation d'énergie. Notion de base connexe Contrôle et régulation Les systèmes vivants maintiennent certains états par des mécanismes de régulation et réagissent aux changements auxquels ils sont soumis.	Sous-domaine 6 Processus de développement – évolution et questions d'avenir Notion de base connexe Variabilité et adaptation Les organismes vivants sont adaptés à leur environnement dans leur structure et leur fonction. L'adaptation est possible grâce à la variabilité. Les bases de la variabilité chez les organismes vivants sont la mutation, la recombinaison et la modification. Notion de base connexe Histoire et parenté Les analogies et la

Source	Degré	Cytologie/ Anatomie/Métabolisme	Traitement de l'information/ Comportement	Immunobiologie	Génétique/ Biologie du développement	Ecologie	Systématique/- Evolution
		<p>dissimilation dans les relations entre les structures cellulaires et l'organisme</p> <p>Catalyseurs moléculaires du métabolisme: enzymologie</p> <p>Notion de base connexe Métabolisme et transformation d'énergie Les organismes vivants sont des systèmes ouverts, qui impliquent des processus métaboliques et de transformation d'énergie.</p>					diversité des organismes vivants sont le résultat de processus d'évolution phylobénétique.
Classification allemande de Ratisbonne (RVK), 2008		<ul style="list-style-type: none"> • Cytologie • Microbiologie • Morphologie et anatomie des plantes • Physiologie des plantes • Morphologie, anatomie et physiologie générales; anatomie, physiologie et biochimie de certains organes et systèmes organiques humains et animaux • Alimentation, digestion, métabolisme, équilibre énergétique, reproduction, physiologie du développement humain et animal, physiologie appliquée 	<ul style="list-style-type: none"> • Le traitement de l'information n'est pas expressément mentionné • Recherche comportementale et psychologie animale 	<ul style="list-style-type: none"> • Immunologie 	<p>Biologie du développement mentionnée dans</p> <ul style="list-style-type: none"> • Alimentation, digestion, métabolisme, équilibre énergétique, reproduction, <u>physiologie du développement humain et animal</u>, physiologie appliquée • Génétique • Biotechnologie 	<ul style="list-style-type: none"> • Ecologie 	<ul style="list-style-type: none"> • Evolution, botanique spécialisée (systématique) • Zoologie spécialisée (taxinomie) • Anthropologie

Les sections ci-après présentent les catégories existantes ou en cours d'élaboration d'une autre dimension, celle des «niveaux d'exigences cognitives». A notre connaissance, il existe trois catégorisations de cette dimension: il s'agit d'une part de la «culture scientifique» de Bybee, et d'autre part à nouveau des standards de formation KMK et des EPA Biologie en Allemagne. En Suisse, il n'existe pas de réflexion analogue pour cette dimension aux degrés secondaires I et II.

4.2.6 Culture scientifique

La proposition de distinguer une dimension de la «culture scientifique» («Scientific Literacy») nous vient des pays anglosaxons. Bybee (1997) propose cinq niveaux de culture scientifique:

Figure III.22: Niveaux de culture scientifique (Scientific Literacy) d'après Bybee (1997)

Niveaux	Caractéristiques
Culture scientifique multidimensionnelle (Multidimensional Scientific Literacy)	Conscience de la signification culturelle des sciences naturelles, capacité de relier des faits et des idées à l'intérieur des disciplines et entre ces dernières, de même qu'avec des problèmes de la société. Compréhension de la dimension philosophique et historique des disciplines.
Culture scientifique conceptuelle et procédurale (Conceptual Scientific Literacy)	Compréhension de la manière dont les différentes composantes d'une discipline s'articulent dans son ensemble et des conséquences que ce lien implique pour les méthodes et les procédures de recherche (exemple: la notion d'évolution englobe celles de la génétique, du métabolisme, de la structure et de la fonction).
Culture scientifique fonctionnelle (Functional Scientific Literacy)	Capacité de lire et d'écrire des textes utilisant un vocabulaire scientifique simple; savoir consistant dans une liste de définitions mémorisables; capacité de reconnaître des relations simples, mais uniquement dans un contexte étroitement limité.
Culture scientifique nominale (Nominal Scientific Literacy)	Capacité d'associer des notions avec des champs de savoir ou des matières; amas de faits appris sans réelle compréhension (théories naïves, concepts erronés).
Absence de culture scientifique (Illiteracy)	Incapacité cognitive de comprendre des questionnements des sciences naturelles et technologiques, ou de relier ces dernières aux sciences naturelles.

4.2.7 Standards de formation KMK

L'échelle des standards de formation KMK ne se réfère pas à la notion de culture scientifique. Elle semble s'inspirer fortement des EPA Biologie et distingue les trois niveaux d'exigences suivants (cf. KMK, 2005, chap. 4.1, p. 16):

Niveau d'exigence I: Reproduction de faits, de méthodes et d'aptitudes.

Ce niveau d'exigence comprend la restitution de connaissances spécifiques et l'application de méthodes et d'aptitudes existantes.

Niveau d'exigence II: Application de faits, de méthodes et d'aptitudes dans un nouveau contexte.

Ce niveau comprend l'application de faits de base connus à un nouveau contexte, différant thématiquement du contexte dans lequel le savoir ou les compétences en question ont été acquis.

Niveau d'exigence III: Formulation de faits nouveaux et réflexion sur ces derniers; application autonome de méthodes et d'aptitudes.

Ce niveau comprend la mise en lumière autonome de faits et de problèmes non connus et la réflexion sur ces derniers sur la base des connaissances préexistantes. Le savoir conceptuel et les compétences sont utilisés notamment pour fonder des explications, des enquêtes, des modèles ou des avis propres.

4.2.8 EPA Biologie

Les EPA Biologie distinguent également trois domaines présentés comme une contribution originale (KMK, 2004, chap. 2.1, p. 13):

Les niveaux d'exigences se différencient principalement par le degré d'autonomie dans le traitement des tâches et par le degré de complexité des processus cognitifs en jeu; ils constituent donc une échelle relative au niveau d'exigence des tâches.

Le niveau d'exigence I comprend

- la disposition de données, de faits, de règles, de formules, de lois mathématiques, etc. propres à un domaine particulier dans leur contexte spécifique appris;
- la description et l'utilisation de techniques de travail et de procédés appris et exercés dans un domaine limité et un contexte récurrent.

Pour la discipline de la biologie, cela comprend

- la reproduction de connaissances de base (connaissance de faits, de relations et de méthodes);
- l'application de méthodes et de modèles dans des exemples comparables;
- l'extraction d'informations dans des textes spécialisés et la réorganisation de ces informations dans des schémas simples (arbres phylogénétiques, organigrammes, etc.);
- l'expression écrite de données, de tableaux, de diagrammes et d'illustrations dans la langue spécialisée;
- la description d'observations macroscopiques et microscopiques;
- la description et le protocole d'expériences;
- la réalisation d'expériences dirigées et l'élaboration de préparations microscopiques;
- l'utilisation compétente de logiciels connus.

Le niveau d'exigence II comprend

- la sélection, l'organisation, le traitement et la représentation autonomes de faits connus selon des critères donnés dans un contexte balisé par des exercices;
- la généralisation autonome de faits appris à des nouvelles questions comparables, à des relations de faits modifiées ou à des procédés transformés.

Pour la discipline de la biologie, cela comprend

- l'application de notions de base dans un nouveau contexte;
- le transfert et l'adaptation de modèles connus;

- la représentation correcte, structurée de façon autonome et adaptée aux tâches de processus biologiques complexes dans le contexte d'exercices;
- la sélection de données, de faits et de méthodes connus pour constituer de nouvelles relations;
- l'extraction ciblée d'informations à partir de matériaux complexes à plusieurs niveaux ou d'une publication scientifique sous un aspect donné;
- la représentation abstraite de phénomènes biologiques comme la représentation et l'interprétation sous forme d'esquisse d'une préparation microscopique non connue;
- l'application d'expériences et de méthodes d'observation connues dans un nouveau contexte;
- l'interprétation de résultats d'études inconnus selon des critères connus;
- le jugement et l'appréciation d'un fait biologique connu;
- la distinction entre représentations communes et connaissances scientifiques.

Le niveau d'exigence III comprend

- le traitement répondant à une démarche planifiée et créative de questions à niveaux multiples dans le but de parvenir de façon autonome à des solutions, des interprétations, des appréciations et des conclusions;
- la sélection consciente et autonome de méthodes et de procédés adéquats appris et leur adaptation à un nouveau contexte.

Dans la discipline de la biologie, cela comprend

- le développement d'une approche personnelle d'un phénomène biologique, par exemple la planification d'une expérience ou d'une expérience imaginaire appropriée;
- le traitement autonome et synthétique de différents matériaux sous l'angle d'un questionnement original propre;
- le développement d'un modèle théorique complexe ou l'adaptation autonome d'un modèle existant;
- le développement d'hypothèses fondées sur la base de différents faits, résultats d'expériences, matériaux et modèles;
- la réflexion de faits biologiques par rapport à la représentation de l'homme;
- le jugement et l'appréciation différenciés et objectifs d'applications biologiques;
- l'argumentation sur la base de données brutes non évidentes: préparation des données, analyse des erreurs et établissement de relations;
- la réflexion critique sur des termes spécialisés de biologie dans le contexte d'informations et d'observations complexes et contradictoires.

4.3 Description de la grille de compétences EVAMAR en biologie

La figure III.23 montre les composants qui ont finalement été retenus pour la grille de connaissances spécifiques en biologie, compte tenu des limitations pragmatiques décrites à la section 4.1. On reconnaît aisément que la dimension des contenus se fonde sur l'analyse des trois ouvrages standard de langue allemande pour l'enseignement de la biologie au niveau gymnasial considérés (degré secondaire II, cf. figure III.18). Ces ouvrages représentent le mieux la division des matières pratiquée dans l'enseignement de la biologie au gymnase. On peut ainsi identifier dix domaines thématiques auxquels tous les manuels standard consacrent un, voire plusieurs chapitres.

Cette classification en dix domaines est cependant encore trop fine pour son utilisation dans le cadre des sous-projets C et D1. Dans son manuel destiné à l'enseignement «Biologie Oberstufe», l'éditeur Cornelsen (2001) a décidé de regrouper certains thèmes fortement imbriqués, comme la génétique et la biologie du développement

ou encore les trois domaines connexes de la neurobiologie, des hormones et du comportement, qui sont traités dans un même chapitre. La classification en sept domaines qui en résulte est cependant encore trop détaillée pour l'utilisation dans le contexte d'EVAMAR II. Pour une question de test (module C) ou un problème de maturité (sous-projet D1) donnés, il est souvent difficile de déterminer si elle se rattache plutôt à la cytologie ou plutôt à la physiologie du métabolisme. Ces deux domaines ont donc été regroupés avec l'anatomie en un seul grand domaine, «Cytologie/Anatomie/Physiologie», sachant que l'anatomie comprend aussi l'enseignement de toutes les formes et structures externes et internes, y compris à très petite échelle (microanatomie). Comme on le voit dans la figure III.21, la classification thématique de la biologie choisie pour EVAMAR II correspond presque entièrement à celle du manuel «Biologie Oberstufe».

En ce qui concerne la dimension des niveaux d'exigences cognitives, le principal critère retenu a été celui de l'adéquation au degré secondaire II. L'échelle de Bybee (1997) est formulée volontairement sans tenir compte des degrés d'éducation et couvre tout le spectre entre l'absence totale de culture et la maîtrise globale du système des sciences naturelles. Elle est de ce fait peu appropriée. Des deux classifications restantes, seule celle des EPA se réfère expressément au degré secondaire II. La fin de l'école moyenne en Allemagne («Mittlerer Schulabschluss») correspond à la fin de l'école obligatoire en Suisse, c'est-à-dire au degré secondaire I. Les standards de formation KMK en biologie ne sont donc pas applicables en l'occurrence. De plus, les trois niveaux d'exigences des EPA Biologie sont formulés plus précisément que ceux des standards de formation KMK en biologie. L'échelle à trois degrés est suffisamment grossière pour permettre de classer sans hésitation les items des questions de test du module C et des questions d'examen analysées dans le sous-projet D1.

Figure III.23: La grille de compétences Connaissances spécifiques en biologie élaborée pour EVAMAR II

		Domaines thématiques (par analogie à des ouvrages standard en allemand pour l'enseignement de la biologie au degré secondaire II)					
		Cytologie/ Anatomie/ Métabolisme	Traitement de l'information/ Comportement	Immuno- biologie	Génétique/ Biologie du développement	Ecologie	Systématique/- Evolution
Niveaux d'exigences cognitives (EPA Biologie)	I						
	II						
	III						

IV Module C: développement et application de tests

1 Généralités et vue d'ensemble

1.1 Objectifs et choix de l'objet des enquêtes

Le module C a été le cadre d'enquêtes auprès des bacheliers et bacheliers. Elles se composaient de tests dans trois disciplines, d'un questionnaire complémentaire et, en Suisse alémanique, d'un test des compétences transversales (TCT).

Il n'est pas possible d'établir une relation directe de cause à effet entre la formation gymnasiale et les résultats chiffrés. Néanmoins, ceux-ci donnent une image des connaissances, des compétences, des savoir-faire et d'autres qualités particulières des élèves peu de temps avant la maturité. Ils permettent de tirer un parallèle avec ce qui pourrait être l'état idéal à l'interface Gymnase-Université, de rechercher des explications utiles aux écarts constatés et aux structures que ceux-ci révèlent, et de formuler des recommandations pour l'avenir du gymnase.

Le questionnaire demandait notamment aux étudiants, en complément à EVAMAR I, d'autoévaluer quelques autres compétences faisant partie de l'aptitude aux études supérieures. Ce point s'est révélé très utile pour deux raisons: d'une part, les données et constats issus des sous-projets A1 et A3 n'offraient guère d'éléments précis pour définir le concept construit de maturité universitaire, et d'autre part les résultats de l'enquête auprès du corps professoral universitaire dans le module demandaient un tel complément. Ce questionnaire a été rempli de manière encore plus complète en Suisse alémanique, en raison du temps imparti.

Il est ressorti des constatations du sous-projet A2 que le test de compétences transversales (TCT) s'imposait (cf. chapitre II, section 4). Le TCT contient quant à lui des items ayant des corrélations, bien attestées par d'autres enquêtes, avec la réussite des études de sciences. Ses items sont validés et proviennent des quatre volets du test d'aptitude aux études de médecine qui ont une valeur prédictive particulièrement haute pour les examens dans les branches d'études universitaires. Ce test additionnel n'a été réalisé qu'en Suisse alémanique, car la direction du projet, au vu de divers signaux reçus, a estimé que l'on ne pouvait plus exiger de consacrer du temps supplémentaire pour ces tests en Suisse romande et au Tessin. Aussi a-t-on renoncé à ce complément dans ces deux régions, étant donné qu'il ne figurait pas dans la planification initiale du projet et qu'il n'en avait pas été fait mention dans les bulletins.

Ces deux démarches de relevé supplémentaires devaient nous permettre de dégager d'autres observations d'une part en relation avec d'autres aspects de l'aptitude aux études universitaires, et d'autre part sur la relation de tests de connaissances standardisés et des compétences plus générales.

Finalement, nous avons procédé aussi au relevé des notes de maturité dans les trois disciplines testées.

1.2 Observations sur la détermination des domaines de test

1.2.1 De la valeur prédictive des tests pour l'aptitude aux études supérieures

Les tests visent à mesurer des aspects de l'aptitude universitaire. Les pages qui suivent développent quelques réflexions sur la valeur prédictive de ces mesures²³.

Comme l'aptitude aux études universitaires se révèle en cours d'études, le moment de la maturité ne permet en principe que de tester des «constructions» ayant montré, selon les recherches empiriques menées jusqu'ici, une valeur prédictive plus ou moins grande en la matière. La littérature portant sur ce domaine, presque exclusivement en langue anglaise, on trouve pour la «maturité universitaire» des concepts construits tels que «academic achievement», «academic aptitude», «academic performance» ou «academic success». Selon les accents mis, les travaux ont analysé divers aspects qui se recoupent, comme

- le niveau général de performance cognitive,
- les connaissances préexistantes,
- les compétences en lecture,
- la personnalité ou
- la motivation.

Le survol bibliographique des pages qui suivent présente une vue d'ensemble concise des résultats de la recherche menée dans ces domaines.

Niveau général de performance cognitive

Il est plausible, et par ailleurs bien attesté dans la littérature, qu'une relation existe entre les capacités cognitives d'une personne et son aptitude à faire des études universitaires (cf. les études cités chez Chamorro-Premuzic, 2007). Cependant, le nombre de points des tests pour le QI ne constitue guère que 25% de la variance de la maturité universitaire (Chamorro-Premuzic & Furnham, 2004; O'Connor & Paunonen, 2007; Neisser et al., 1996), en général même bien moins. On attribue cela au fait que les étudiants disposent en principe déjà d'une capacité de performance cognitive plus homogène que l'ensemble de la population (Furnham, Chamorro-Premuzic & McDougall, 2003). Il est intéressant de noter, en relation avec les facteurs motivationnels, que la stratégie d'étude du *Deep Processing* présente une corrélation de ($r = .32$) avec le QI mesuré dans le «Wonderlic Personnel Test», et que ce dernier à son tour a une corrélation de ($r = .24$) avec les notes des examens universitaires (Chamorro-Premuzic & Furnham, 2008).

Dans EVAMAR II, les compétences cognitives relativement indépendantes du savoir sont mesurées par le test de compétences transversales (TCT).

Connaissances préexistantes

Lors de l'écoute et de la lecture d'une information nouvelle, les connaissances préalables favorisent la profondeur de la compréhension, la quantité des informations rappelées ensuite et la correction de ce qui a été compris. Comme l'attestent les recherches empiriques, la qualité des connaissances préexistantes détermine dans une large mesure le degré de finesse et d'achèvement de la mise en relation du contenu de ce qu'on vient de lire avec ce savoir antérieur; de ces connaissances préexistantes dépendent aussi l'opportunité de fait de cette connexion et la qualité de la mise en mémoire des nouveaux contenus (par ex. B. Bartlett, 1932; Ausubel, 1968; Bloom, 1976; Chiesi, Spilich & Voss, 1979; Wilson & Anderson, 1986; Recht & Leslie, 1988; Alexander et al., 1989; Dochy, 1992; Tobias, 1994; Kim & Van Dusen, 1998; Thompson & Zamboanga, 2004; Shapiro, 2004; Baumert et al., 2007; Hailikari, Nevgi & Komulainen, 2008; Ozuru, Dempsey, McNamara, 2009). Selon l'étude de synthèse de Dochy, Segers & Buehl (1999), 95% des 183 publications sélectionnées ont pu prouver l'effet favorable que des

²³ Des remarques sur la valeur prédictive des notes de maturité se trouvent au chapitre I, section 1.

connaissances préexistantes dans le domaine concerné ont sur les performances d'apprentissage. Spécialement lorsque le texte du cours présente une logique interne peu affirmée (cohésion limitée), les lecteurs et lectrices disposant de connaissances préalables importantes peuvent plus aisément, pendant la lecture, tirer des conclusions favorables à l'apprentissage sur les relations locales et globales entre les diverses phrases. Cela leur permet de maintenir une représentation du texte plus cohérente et finalement de la stocker dans la mémoire (McNamara & Kintsch, 1996; McNamara, 2001). Un texte présente une part de cohésion limitée dans les cas suivants quand les conditions suivantes sont réunies:

- de nombreuses notions, concepts et théories importantes pour la compréhension sont supposés connus comme savoir d'arrière-plan, et les affirmations du textes ne sont pas explicitement mises en relation avec les notions, concepts et théories en question;
- le texte présente une faible redondance argumentative;
- le texte comporte des sauts dans la pensée et des imprécisions de fond;
- le texte ne contient guère d'éléments d'aide structurelle aidant à la compréhension comme des sous-titres, mises en évidence, encadrés (Best, Rowe, Ozuru & McNamara, 2005).

Un bagage de savoir préalable riche et bien structuré revêt donc une importance toute particulière dans le contexte universitaire, qui présentera fréquemment à l'étudiant des cours d'une cohésion limitée, soit parce que les connaissances préalables doivent être présupposées (pour des raisons d'économie de la formation) soit aussi parce qu'il s'agit de documents susceptibles d'être améliorés (par ex. Beck, McKeown, Sinatra & Loxterman, 1991; Britton, Gulgoz & Glynn, 1993). Les connaissances préalables n'ont cependant pas toujours un effet favorable à la profondeur de compréhension du texte, car des conceptions fausses qu'elles contiennent peuvent venir parasiter les idées développées dans le document à lire (par ex. B. West & Pines, 1985; Diakidoy, 1999), au point de compromettre même la réussite de l'apprentissage (Alexander & Judy, 1988; Vosniadou & Brewer, 1994; Kendeou & van den Broek, 2007; Hailikari, Nevgi & Komulainen, 2008). Il est intéressant de relever que les conceptions erronées n'entravent pas l'apprentissage parce qu'elles ralentiraient le processus de lecture, mais bien plus parce que les lecteurs chez qui elles existent ont des exigences limitées en matière de profondeur de compréhension, donc de standards de cohérence (cf. plus bas les explications sur le *Landscape model*) (van den Broek, Ridsen & Husebye-Hartmann, 1995; Kendeou & van den Broek, 2007); ces lecteurs ne tentent donc pas de tirer au clair les contradictions entre leur savoir préexistant et le cours qui est devant leurs yeux par le biais d'efforts supplémentaires (par ex. retour à une partie du texte déjà lue, autres activations de la mémoire, recours à des moyens auxiliaires externes ou à Internet). Les lacunes dans l'apprentissage n'apparaîtront dans ces cas-là que lors d'un test de performances (Kendeou & van den Broek, 2007).

Cet aspect de l'analyse – le degré d'élaboration des connaissances préalables, y compris des conceptions erronées (latentes) – a fait l'objet des tests de connaissances et de compétences dans les disciplines de mathématiques et de biologie.

Compétences en lecture (habileté à lire, stratégies et techniques de lecture)

Des compétences de lecture développées et des techniques à la hauteur peuvent dans une certaine mesure compenser un savoir préalable lacunaire (Garner, 1987; Voss & Silfies, 1996; O'Reilly & McNamara, 2007). S'agissant de textes scientifiques, la compétence de lecture se définit essentiellement par l'aptitude à reconnaître des liens entre les phrases et entre les paragraphes (lecture référencée, «inference making», cf. la bibliographie chez O'Reilly & McNamara, 2007). Les techniques de lecture qui s'y associent englobent des compétences telles que «savoir dégager d'abord une vue d'ensemble», «résumer les affirmations centrales ou les reformuler avec ses propres mots», «lire plusieurs fois les passages difficiles», «rechercher des renseignements en avant ou en arrière dans le texte», «lire à haute voix», «faire à chaque instant le bilan critique de la compréhension», «recourir à des techniques de mise en évidence», «utiliser des manuels», etc. (Mokhtari & Reichard, 2002).

Dans le volet d'EVAMAR II qui teste des compétences langagières en langue première, la plupart des items se rapportent à la compétence de lecture (compréhension de l'écrit).

Connaissances préalables et capacités de lecture dans le contexte

Il y a des années déjà que la recherche, tant théorique qu'empirique, analyse les processus de compréhension qui ont cours lors de la lecture et la manière dont les connaissances préalables influent sur la représentation mémorielle d'un texte. L'une des synthèses les plus remarquées des monographies sur le sujet est le «Landscape Model» (van den Broek, Risdén, Fletcher & Thurlow, 1996; van den Broek, Young, Tzeng & Linderholm, 1999; van den Broek, Rapp & Kendeou, 2005) qui présente des similitudes avec le modèle «construction-intégration» «Construction-Integration Model» (Kintsch, 1988, 1998) et d'autres modèles articulés en deux phases (notamment Long, Seely & Oppy, 1996; Richards & Singer, 2001). Dans l'intervalle, les postulats de base du «Landscape Model» ont été implémentés en un modèle informatique. Les prédictions tirées de parcours de lecture simulés peuvent ainsi être mises en parallèle avec les données empiriques, permettant ainsi une validation plus fine du modèle (van den Broek, Virtue, Everson, Tzeng & Sung, 2002; van den Broek, Kendeou, Sung & Chen, 2003; Tzeng, van den Broek, Kendeou & Lee, 2005).

Le mérite spécifique du «Landscape Model» consiste à relier les processus de récupération en mémoire et les mécanismes de construction rapide de renseignements nouveaux et à représenter leur interaction sous forme de modèle jusque dans le détail du plan des cycles de lecture concrets.

Si l'on envisage l'élaboration du texte selon le postulat fondé sur la mémoire, l'information que l'on vient de lire (grossie de l'information qui vient d'être activée dans la mémoire de travail) déclenche l'activation d'une autre information associée, stockée dans la mémoire à long terme. Plus l'association entre l'information qui vient d'être lue et l'information de la mémoire à long terme est élaborée, plus aisée, plus rapide et plus solide sera l'activation de l'information à long terme par l'information que l'on vient de lire. L'activation de l'information peut se produire aussi au sein de la représentation en train de se faire sur la base d'un épisode du texte lu jusque-là (le modèle discursif). Cette représentation ne se trouve pas encore stockée dans la mémoire à long terme, mais elle ne se trouve pas non plus dans la mémoire de travail. Les processus fondés sur la mémoire sont considérés comme «gratuits», dans la mesure où ils sont passifs et se déroulent de manière autonome, sans être soumis à une stratégie consciente du lecteur.

Si l'on considère l'élaboration du texte selon le postulat de la construction, les lecteurs tentent, quand ils commencent à lire, d'atteindre consciemment des buts implicites et explicites ou des standards en relation avec la construction du sens ou d'accéder à une certaine profondeur de compréhension. Ils parviennent à atteindre ces standards de cohérence (*standards of coherence*) par le rappel d'informations qu'ils tirent du passage du texte déjà lu (au besoin en relisant), de la représentation du texte à lire qui se constitue petit à petit dans leur esprit (c.-à-d. dans leur mémoire de travail), ou de leur mémoire à long terme. En cas de discordance entre les diverses informations collectées, le lecteur au bénéfice de standards de cohérence élevés met en œuvre d'autres procédures constructives pour tenter de résoudre la contradiction (van den Broek & Kendeou, 2008). Des connaissances préexistantes inappropriées ou des conceptions erronées sont donc de nature à entraver sérieusement la fluidité de la lecture si le lecteur obéit à des standards de cohérence élevés. L'intervention des critères de cohérence en relation avec la profondeur de lecture visée va dépendre du but de la lecture: s'agit-il d'une lecture de divertissement ou d'étude? Le niveau desdits standards de cohérence détermine pour une large part la proportion de procédures constructives (constructivistes) engagées en cours de lecture.

Le Landscape Model fait la synthèse de ces deux postulats. Il part du principe que les deux processus interagissent de façon dynamique, exploitent en alternance les résultats de l'autre, et pourtant qu'ils peuvent entrer en conflit. Cette interaction, Van den Broek et ses collaborateurs se la représentent de la manière suivante: le lecteur avance dans une suite de «cycles de lecture» - il faut entendre par là la saisie d'une phrase environ, ou en général une unité d'information. A chaque cycle, le lecteur évalue, selon les standards de cohérence qu'il a choisis, le résultat d'une activation de connaissance de savoir passive qu'une portion de texte a mise en œuvre. Si l'information activée ne suffit pas à comprendre le texte à la profondeur voulue, le lecteur entreprend d'autres efforts de construction (retour en arrière dans le texte, autres activations mémorielles, consultation de moyens auxiliaires comme un dictionnaire ou Internet). Si les standards de cohérence sont bas,

les efforts de construction ne se font pas, ils n'ont pas lieu non plus si l'information activée ne suffit pas à faire naître une présentation du texte cohérente.

Facteurs de la personnalité

Une étude de synthèse publiée récemment et qui recense une vingtaine d'études choisies arrive à la conclusion que le caractère consciencieux (*conscientiousness*) est de loin le paramètre le plus fort et le plus pertinent pour déterminer l'aptitude aux études universitaires (*academic performance*, r moyen = 0.24, 90%-intervalle de confiance entre $r = 0.12$ et $r = 0.36$) (O'Connor & Paunonen, 2007). Cette dimension du caractère consciencieux constitue, avec les variables du névrosisme (*neuroticism*), de l'extraversion (*Extraversion*), de l'ouverture à l'expérience (*openness to experience*) et de l'agréabilité (*agreeableness*), le modèle en cinq facteurs utilisé dans la psychologie de la personnalité (McCrae & Costa, 1997). Des sous-facteurs du caractère consciencieux que sont la volonté de performance (*achievement-striving*), l'autodiscipline (*self-discipline*) et le sens du devoir (*dutifulness*) présentent une corrélation encore meilleure avec la maturité universitaires que le facteur hyper-ordonné de la conscience. Cette recherche intense de performance englobe les traits de l'individu ambitieux, persévérant, appliqué, soigneux et endurant. Quant au sous-facteur de d'autodiscipline, il comporte les caractéristiques de la motivation, de la capacité de mener les choses à leurs terme, et l'aptitude à résister aux diversions. Le sous-facteur du sens du devoir définit le besoin qu'éprouve le sujet de tenir ses engagements moraux (O'Connor & Paunonen, 2007).

Aux termes de l'article 5 du RRM 95 (1995), les gymnases ont pour mission, entre autres, de développer la volonté des élèves; cela indique que ces trois sous-facteurs du caractère consciencieux font partie des qualités ciblées par la formation que dispensent les écoles de maturité. Dans le cadre d'EVAMAR II, c'est le questionnaire complémentaire qui a permis de relever des données relatives à la volonté de performance et à l'autodiscipline.

Facteurs motivationnels

Une excellente vue d'ensemble est contenue dans l'étude – la plus développée à ce jour – où Fenollar, Roman & Cuestas (2007) présentent de façon synthétique en un modèle conceptuel intégratif des facteurs de motivation de l'aptitude aux études universitaires bien attestés empiriquement. Selon leur idée centrale, les stratégies du traitement de l'information relevées (en surface vs en profondeur) et de la gestion de l'énergie («effort») prennent forme sous l'effet de certains objectifs de performance et le degré d'efficacité personnelle subjectivement ressenti (*self-efficacy*). Ces objectifs de performance d'une part influent directement sur l'aptitude aux études universitaires, et d'autre part ont des retombées sur la nature des stratégies du traitement de l'information. Fenollar, Roman & Cuestas (2007) réunissent dans leur modèle quatre objectifs de performance relativement indépendants qui trouve une place toujours mieux établie dans la littérature (Ames & Archer, 1988 et Elliot & Thrash, 2001):

- les buts de maîtrise (*Mastery Goals*): on entend développer ses compétences, augmenter ses connaissances et, par le biais d'un apprentissage soutenu, accéder à une compréhension et une maîtrise de la matière;
- les buts d'approche de la performance (*Performance-approach Goals*): on veut donner l'image de l'étudiant qui réussit à s'imposer et qui l'emporte sur les autres;
- les buts d'évitement de la performance (*Performance avoidance Goal*): on a le souci de ne pas laisser paraître son incompetence, on va cacher ce qui pourrait la révéler (aux dépens d'un apprentissage de longue durée);
- les buts d'évitement de l'effort (*Work Avoidance Goals*): ici, l'étudiant veut liquider le travail avec le moins d'effort possible et éviter les tâches qui le sollicitent particulièrement.

La mesure de sa propre efficacité englobe les représentations intérieures (subjectives) par lesquelles le sujet se convainc de sa capacité d'effectuer avec succès une action requise pour atteindre le but fixé (Bandura 1977).

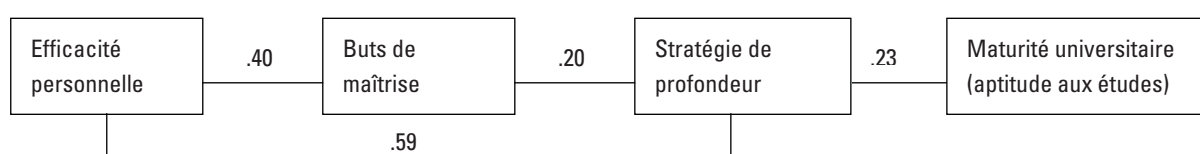
La différenciation d'un traitement approfondi ou superficiel de l'information a été inaugurée par Marton & Säljö (1976), puis a trouvé sa place dans la littérature touchant à la psychologie cognitive grâce à Tait & Entwistle (1996).

Les étudiants qui suivent une stratégie de profondeur (**Deep processing**) s'efforcent de comprendre la signification de ce qu'ils apprennent en élaborant des relations entre les notions, en jetant des ponts entre les connaissances préexistantes et la matière à apprendre et en adoptant une position critique.

Les étudiants qui adoptent une stratégie de surface (**Surface processing**) concentrent leurs efforts sur des stratégies de mémorisation et se dispensent de tenter de comprendre la signification de ce qu'ils «apprennent».

Fenollar, Roman & Cuestas (2007) ont été en mesure de confirmer l'hypothèse selon laquelle les stratégies d'apprentissage ont un effet dynamique entre les objectifs d'apprentissage et l'aptitude aux études universitaires. Ils ont dégagé des corrélations hautement significatives ($p < 0.01$) entre les couples de paramètres suivants, reliés à chaque fois par un segment de droite (cf. figure IV.1):

Figure IV.1: Stratégies d'apprentissage, objectifs de performance et succès dans les études



Les relations positives entre ces quatre paramètres trouvent leur confirmation dans un grand nombre d'autres analyses (on en trouvera les titres chez Fenollar, Roman & Cuestas, 2007; cf. ci-dessus les explications données à ce sujet dans les paragraphes consacrés au «Landscape Model» de la compréhension de l'écrit).

Il va sans dire que dans ce domaine aussi une mission du gymnase consiste à offrir un bon enseignement aux élèves et à former ainsi des bacheliers et bachelères assez confiants pour savoir résoudre eux-mêmes des questions difficiles (*Self-efficacy*), assez exigeants pour vouloir véritablement comprendre la matière qu'ils étudient, sans se contenter d'un apprentissage mémorisé ou d'une réussite de surface (*Mastery Goals* et *Deep Processing*).

Les travaux d'EVAMAR II ont analysé en partie dans quelle mesure ces objectifs de l'aptitude aux études universitaires sont atteints.

Le tableau IV.1 ci-dessous présente une récapitulation des paramètres les plus importants d'un haut degré d'aptitude aux études universitaires. Ces qualités peuvent et doivent toutes trouver, dans le cadre de la formation gymnasiale, un bon terrain de développement dans l'esprit de l'article 5 du RRM 1995. On observe que les enquêtes d'EVAMAR II touchent à une large palette de facteurs cognitifs (niveau général de compétence cognitive, connaissances préexistantes et compétences langagières) et, au-delà, à des éléments importants des facteurs de la personnalité: il en résulte une image riche et complexe des composantes de la maturité universitaire.

Tableau IV.1: Composantes importantes d'une maturité universitaire élevée

Aspects	Composantes forts de l'aptitude générale aux études	Etudes de synthèse importantes	Nature de l'analyse dans EVAMAR II
Niveau général de compétence cognitive	<ul style="list-style-type: none"> compétences cognitives générales / intelligence 	Chamorro-Premuzic (2007)	Items du test de compétences transversales (TCT)
Connaissances préexistantes	<ul style="list-style-type: none"> degré d'élaboration du savoir préalable nombre et pertinence des représentations erronées 	Dochy, Segers & Buehl (1999)	Tests (de connaissances et de compétences) en mathématiques et en biologie
Compétences en lecture	<ul style="list-style-type: none"> habileté à lire richesse du répertoire des stratégies et des techniques de lecture utilisées 	O'Reilly & McNamara (2007)	Test des compétences langagières en langue première, en particulier des techniques de lecture
Facteurs de la personnalité	<ul style="list-style-type: none"> recherche intense de performance autodiscipline sens du devoir 	O'Connor & Paunonen (2007)	Facteurs saisis en partie par le biais de quelques items du questionnaire complémentaire
Facteurs motivationnels	<ul style="list-style-type: none"> compétence subjective (Self-efficacy): s'estimer capable de résoudre des problèmes, même difficiles buts de maîtrise et cheminement en profondeur (Mastery Goals, Deep Processing): éprouver le besoin de comprendre véritablement la matière du cours et ne pas se contenter d'un apprentissage par cœur ou d'une réussite superficielle 	Fenollar, Roman & Cuestas (2007)	Facteurs saisis en partie par le biais de quelques items du questionnaire complémentaire

1.2.2 Choix de la biologie comme branche des sciences naturelles

Ainsi que le décrit le chapitre I, EVAMAR II avait reçu mandat devait développer et appliquer des tests dans trois disciplines seulement, langue première, mathématiques et une discipline des sciences expérimentales – non encore définie au moment de l'attribution du mandat. En décembre 2006, le groupe de pilotage, à la lumière des premiers résultats du module A, a fait son choix parmi les trois disciplines envisagées, soit biologie, chimie, physique et a opté pour la première nommée. Les motifs essentiels qui ont motivé ce choix sont les suivants: Dans le module C, il s'agit, en ce qui concerne les sciences naturelles, de mesurer l'aptitude des futurs diplômés à montrer dans un test écrit portant sur le savoir et le savoir-faire, ceci sans moyens auxiliaires, qu'ils se souviennent assez exactement de faits, concepts et procédés généraux importants pour les études et qu'ils les ont compris assez exactement. Dans ce contexte, les faits, concepts et procédés considérés comme «généralement importants pour les études» sont ceux qui sont censés être effectivement connus et assimilés selon les matériaux d'enseignement et les examens intermédiaires du premier semestre de *toutes* les voies d'étude.

La question de savoir si les faits, concepts et procédés effectivement testés devaient être en premier lieu de nature biologique, chimique ou physique a été résolue sur la base de deux réflexions:

1. Puisqu'en dernier ressort il s'agissait toujours d'examiner l'«aptitude générale aux études», il a fallu décider dans laquelle de ces trois branches sont enseignés avec le plus de probabilité, et ceci jusqu'à la maturité, les faits, concepts ou procédés qui sont censés être connus au début des études dans des directions d'études les plus diverses possibles.
2. Pour pouvoir exploiter au maximum les résultats, il a en outre fallu décider dans laquelle des branches biologie, chimie ou physique sont enseignés avec le plus de probabilité, et ceci jusqu'à la maturité, les faits, concepts et procédés qui ont le moins été évalués par le biais des compétences déjà testées dans le cadre du test de mathématiques.

La première réflexion a mené à une réponse claire en faveur de la biologie. Contrairement à la physique et à la chimie, les sciences techniques des sciences naturelles, mais aussi les sciences humaines et les sciences sociales ont recours à des concepts biologiques. On pourrait mentionner à ce propos la prise en compte du point de vue évolutionnaire dans les sciences économiques ou la psychologie. D'autre part, selon les résultats du module A, la chimie et la physique jouent un rôle moins important dans les sciences humaines et sociales. De toutes les sciences naturelles, la biologie semble avoir le plus de facilité à franchir la «frontière culturelle» entre les sciences humaines et les sciences tout court (voir à ce propos Snow, 1963).

La deuxième réflexion a mené à une réponse défavorable à la physique. En effet, si les faits, concepts et procédés enseignés en physique et d'importance pour les études se différencient de ceux enseignés en mathématiques, certains procédés, mais aussi quelques concepts mathématiques sont malgré tout nécessaires pour résoudre un problème de physique typique. Au niveau gymnasial surtout, les problèmes de physique doivent souvent tout d'abord être impérativement traduits dans la langue mathématique des équations, pour ensuite être résolus à nouveau avec des méthodes mathématiques. Par ailleurs, le fait que les notes scolaires gymnasiales en physique sont plus proches de celles obtenues en mathématiques que de celles obtenues en biologie/chimie indique que les exigences cognitives des mathématiques et de la physique sont apparentées (Heller, 2002).

La première et la deuxième réflexion justifiaient ensemble le choix de la biologie. En ce qui concerne la chimie, la deuxième réflexion ne menait pas à un choix négatif à l'inverse de la physique, mais la première réflexion ne lui était pas non plus favorable. Le choix s'est donc porté sur la biologie, et les faits, concepts et procédés relevant des sciences naturelles testés dans le cadre d'EVAMAR II ont été en premier lieu de nature biologique..

2 Description des instruments de relevé

2.1 Démarche suivie pour la construction des tests: remarques générales

La construction des items du test s'est fondée sur les grilles de compétence établies dans le module B, sur les éléments de connaissance et de compétence saisis dans la base de données résultant des sous-projets A1 et A3 et sur le matériel de cours concret en rapport avec le sujet. Nous avons veillé à obtenir une répartition aussi représentative que possible entre les champs de compétence d'une part et entre les éléments de connaissance et de compétence identifiés d'autre part. A l'exception, déjà mentionnée, du domaine de la langue première, il n'a pas été nécessaire de corriger cette base à la lumière des résultats de l'enquête auprès du corps professoral réalisée dans le sous-projet A4.

Les analyses du contenu du matériel de cours (sous-projet A1) a permis d'identifier un grand nombre d'expressions techniques, de classements, de catégories, de principes, de théories et de modèles qu'il faut considérer comme des connaissances de départ étant donné qu'ils ne font plus l'objet d'explications dans la documentation du cours, mais qu'ils sont présumés acquis à l'entrée des études.

Chaque item de test se réfère à un locus d'un texte du matériel de cours (sous-projet A1). Chaque item a été construit dans le but d'évaluer si une bachelière ou un bachelier a compris une unité de sens avec une certaine profondeur. Les unités de sens visées peuvent être des termes spécialisés, des catégories, des classifications, des principes et même une théorie complète ou un modèle complet (cf. chapitre II). Ce qui est déterminant, c'est que le matériel de cours concerné ne renseigne pas sur la compréhension de l'unité de sens, mais la présume acquise. Les items permettent donc de mesurer l'étendue et la qualité des connaissances de départ (prérequis) et de leur compréhension. Nous sommes partis des hypothèses de travail suivantes:

1. Plus la qualité des connaissances de départ est élevée, meilleure est la compréhension du texte spécialisé (cf. section 1.2.1). Plus les réponses correctes aux items sont nombreuses, plus aisée sera la compréhension du matériel de cours pour les passages où apparaissent les termes spécialisés, classifications, catégories, principes, généralisations, théories ou modèles concernés.
2. Plus la qualité de la compréhension spécifique du texte est élevée, meilleure est l'aptitude aux études universitaires dans la discipline qui emploie de telles unités de sens.
3. Plus la qualité de la compréhension spécifique du texte est élevée et plus grande est l'aisance de l'étudiant à rappeler ces unités de sens «scientifiques» de la mémoire à long terme dans le contexte de n'importe quelle discipline d'étude, plus l'aptitude générale aux études universitaires est forte.

Outre les principes généraux qui régissent la construction d'items de tests, une autre nécessité s'imposait à EVAMAR II: pour analyser la «maturité universitaire» à la lumière des exigences concrètes des études, il fallait que les items reflètent aussi fidèlement que possible les exigences cognitives qui attendent réellement les étudiantes et étudiants dès le premier semestre. Tous les items se réfèrent à la situation concrète de lecture et de compréhension de documents de cours (photocopiés et manuels). Pour cette raison, le corps de la question et les options de réponse intègrent souvent les unités de sens dans un contexte de détail le plus semblable possible à celui où elles apparaissent effectivement. Ce contexte rend la plupart du temps les items plus difficiles à résoudre que si l'unité de sens se trouvait présentée isolément.

Les items du test de compétences transversales ont été mis à disposition par le Centre pour le développement de tests et le diagnostic (CTD) de Fribourg.

Afin de pouvoir intégrer toute l'étendue des grilles de compétence, nous avons conçu une somme d'items dont la résolution complète aurait pris à l'élève 720 minutes. Le temps imparti à l'enquête n'était pas suffisant pour soumettre à chacun la totalité des items. Nous disposions seulement de 45 minutes par test. Pour garantir une évaluation fiable des aptitudes des bachelières et bacheliers (paramètre des personnes), mais aussi de la difficulté des items (paramètre de difficulté), nous avons choisi une configuration tournante dans laquelle les items étaient reliés entre eux («Rasch Common Item Equating», soit une matrice multicritères reliant des blocs d'items similaires de manière que les connaissances et compétences relevées chez la personne ne se rapportent pas uniquement au cahier de test qu'elle a rempli) (Wright & Stone, 1979). De cette manière, nous avons dû préparer sept cahiers pour la biologie, six pour les mathématiques, cinq pour la langue première «Allemand», quatre pour chacune des langues premières «Italien» et «Français» et quatre cahiers également pour les compétences transversales. Les cahiers ont été distribués au hasard aux élèves, en veillant à ce que chaque item soit traité par un même nombre de bachelières et bacheliers. Le tableau IV.2 indique la répartition des blocs sur les cahiers avec l'exemple de la biologie.

Tableau IV.2: Les sept cahiers de tests en biologie

	Cahier 1	Cahier 2	Cahier 3	Cahier 4	Cahier 5	Cahier 6	Cahier 7
Demi-cahier A	BIO 1	BIO 2	BIO 3	BIO 4	BIO 5	BIO 6	BIO 7
Demi-cahier B	BIO 2	BIO 3	BIO 4	BIO 5	BIO 6	BIO 7	BIO 1

Les critères déterminants lors du choix des types d'items étaient, à l'instar d'enquêtes de taille comparable (TIMSS, PISA), ceux de l'objectivité et de l'économicité de l'analyse. Nous avons retenu le moins possible d'items à réponse ouverte et circonstanciée (la plupart en langue première) car ils exigent de lourds travaux de codification pour assurer l'objectivité de la correction. Il n'est cependant pas possible de renoncer à ce type d'items, en particulier pour la langue première, dans une analyse qui vise à mesurer les compétences dans toute leur étendue. Notre objectif n'en a pas moins été de garantir l'objectivité la meilleure possible, grâce à une formation soigneuse des correcteurs et correctrices, entre autres mesures. En nous limitant à la forme écrite, nous avons restreint le nombre de champs de compétence testables pour la langue première (par ex. pas de compétences orales ou auditives).

2.2 Description de l'évaluation qualitative des experts et des prétests (y compris procédure de révision et modification des items)

La première version des items élaborés a été soumise à une procédure d'examen de la qualité et de révision selon les étapes suivantes:

- évaluation des items par des experts immédiatement après la construction du test, et première révision
- passation des tests auprès de 100 à 180 bacheliers et bacheliers faisant partie des classes terminales des gymnases choisis; contrôle et révision éventuelle des items sur la base des critères appliqués dans le modèle de Rasch (passation «Infit») indice de discrimination des items, répartition équilibrée des distracteurs et allure conforme au modèle de la courbe caractéristique des items «Item Characteristic Curve» (ICC) (cf. les définitions dans la section 2.4).

Le choix des écoles s'est fait d'une manière pragmatique (critères de l'accessibilité, disposition à la collaboration). Les prétests ont été effectués fin février/début mars 2007.

L'évaluation qualitative des items de la langue première «Allemand» a été assumée par un expert de la Haute Ecole pédagogique de Zurich. Une première passation d'une version qui ne comportait pas encore d'items destinés à mesurer la compétence grammaticale a eu lieu en septembre 2006 à l'école cantonale de Zoug, avec 65 bacheliers et bacheliers, puis les items ont déjà été révisés une première fois. Le deuxième prétest, qui comportait tous les items, a eu lieu juste après la validation des experts aux Ecoles cantonales de Sargans et de Frauenfeld, rassemblant environ 180 participants. Les items traduits en français et en italien ont alors été spécialement passés en prétest. Tous les items de test ont été soumis à une autre optimisation commune après les prétests effectués dans les trois régions linguistiques (cf. chapitre I, section 5.2).

L'évaluation qualitative des items en mathématiques a été effectuée par des experts de l'EPFZ. Au vu des délais très rapprochés, tous les items n'étaient pas encore prêts pour cette première étape de validation. C'est pourquoi certains ont fait l'objet d'une autre expertise dans le cadre de la validation qualitative de la version traduite en français. Le prétest a été passé à l'ancienne Ecole cantonale d'Aarau dans huit classes terminales comprenant en tout 163 bacheliers et bacheliers. Après cela, le lot des items de test a été remanié sur les plans rédactionnel et spécifique et enrichi de nouveaux items.

Le prétest en biologie s'est déroulé au Freies Gymnasium Zurich avec la participation de 110 élèves.

2.3 Les traductions dans les deux autres langues nationales: généralités pour les trois disciplines testées

La traduction des items dans les deux autres langues nationales (italien et français) est une démarche très complexe pour diverses raisons. Deux variantes possibles étaient en discussion:

- Variante 1: traduction des items allemands dans l'une des deux autres langues, puis rétroversion en allemand par une autre personne. En cas de divergence, on procède à une correction de la traduction.
- Variante 2: double traduction des items par deux personnes, comparaison des résultats et discussion et aménagements éventuels en cas de divergence.

Il ressort des expériences faites lors d'analyses internationales comme PISA et TIMSS que la variante 1 est notablement plus coûteuse à cause des divergences qui se manifestent souvent. Même l'engagement de traducteurs de haut niveau ne permet pas d'éviter ces écarts. Pour des considérations d'ordre financier, EVAMAR II a donc opté en règle générale pour la variante 2.

2.4 Démarche suivie pour la détermination des items intégrés dans l'analyse générale

Pour être retenus dans l'analyse générale, les items devaient répondre aux exigences statistiques minimales posées aux tests; ils devaient, en particulier,

- être compatibles avec le modèle de Rasch, apparaît dans l'Infit: cet indice d'ajustement calculé pour chaque item permet de connaître indique le nombre de réponses inattendues qui peuvent apparaître dans le cadre du modèle de Rasch. Dans cette hypothèse, l'Infit a une espérance mathématique de 1. Si un item a une valeur d'Infit ayant un écart statistiquement significatif de 1, l'item en question n'est pas compatible avec le modèle de Rasch et doit être écarté. La valeur de l'Infit devrait se situer entre 0.8 et 1.2.
- présenter un indice minimal de discrimination de 0.2 (pour la justification, cf. par exemple Bühner, 2006, p. 140; Eberle, 1986). On entend par discrimination le coefficient de corrélation (point bisériel) de l'item et du score total non corrigé du test entier pour chaque sujet. Un indice de discrimination élevé signifie par exemple que les bacheliers et bacheliers obtenant un score général élevé vont résoudre correctement l'item, mais que ceux qui ont des valeurs générales basses vont donner des réponses erronées à cet item.
- avoir une courbe caractéristique («Item Characteristic Curve», courbe caractéristique d'item, CCI) conforme au modèle de Rasch, c'est à dire stable.
- faire l'objet d'un examen complémentaire si l'on observe une faible homogénéité des répondants; dans un tel cas distinguant les régions linguistiques, il fallait exclure tout lien causal avec les difficultés de traduction. La vérification de la traduction s'est fondée sur le calcul la fonction différentielle d'item (DIF «Differential Item Function»); en d'autres termes, on a tenté de déceler une éventuelle interaction entre un item de test et la région linguistique. On parle d'une fonction différentielle quand certains items déterminés présentent des probabilités différentes d'être résolus correctement pour des répondants aux compétences similaires, mais de régions linguistiques différentes. L'analyse se complique si l'on considère que les fonctions différentielles d'un item peuvent tirer leur origine autant de problèmes de traduction que de plans d'études dissemblables. Les critères appliqués à cet examen étaient plus ou moins stricts selon les disciplines:
- En langue première, la différence entre les régions alémanique et romande ne devait pas dépasser la valeur de 0,5 Logit sur l'échelle des difficultés d'items. Un item au-dessus de cette valeur, l'item était malgré tout maintenu s'il manifestait une différence importante également entre les deux grandes régions alémanique de Berne et de Zurich (en dehors du seuil de signification de 99%). Ce dernier paramètre peut indiquer que la différence n'est pas due à la traduction, mais à d'autres causes. Ce sont précisément de telles variations que l'analyse doit révéler.
- En mathématiques, la différence entre les régions linguistiques alémanique et romande devait se trouver à l'intérieur du seuil de signification de 99%); si cette limite était dépassée, on maintenait malgré tout un item pour lequel la différence entre les deux grandes régions alémaniques de Berne et de Zurich était aussi considérable (en dehors du seuil de signification de 99%). Comme plus haut, on pouvait dans un tel

cas de figure induire d'autres causes que la traduction. Comme ci-dessus, c'est la mission d'EVAMAR II de mettre en évidence de telles discordances.

- En biologie, la différences entre les régions alémanique et romande devait se trouver d'abord à l'intérieur du seuil de signification de 99%); comme la stricte application des premiers critères ne laissait subsister que 56 items au total, le critère de différenciation linguistique a été quelque peu assoupli et adapté à l'usage fait pour la langue maternelle: la différence entre les deux régions linguistiques ne devait pas dépasser la valeur de 0,5 Logit sur l'échelle des difficultés d'items. Il est possible de justifier aussi cette application moins stricte du critère par le fait que les différences dans les degrés de difficulté reflètent des variations de programmes, elles-mêmes explicables par le fait que la formation et le perfectionnement des maîtres se déroulent de façon indépendante selon les régions linguistiques: on pouvait donc exclure un simple lien causal des éventuelles erreurs de traduction. Le critère de la différence entre les «grandes régions alémaniques Zurich et Berne» n'a pas permis de récupérer d'autres items écartés. L'analyse générale a été calculée et présentée ensuite sur la base de la variante de choix assouplie. Dans les cas où les calculs de contrôle effectués en parallèle donnaient lieu à des différences par rapport à la variante stricte, il en était rendu compte et on procédait à l'analyse voulue.

Cette procédure a finalement permis de confirmer 124 items comme base d'analyse générale en langue première, 117 items en mathématiques et 56 (calcul strict) ou 91 items (variante assouplie) pour la biologie. Le test de compétences transversales a été fondé sur les 42 items qui ont tous pu être intégrés à l'analyse générale.

2.5 Le test de langue première

2.5.1 Précisions sur l'élaboration du test

2.5.1.1 A propos de la construction des items

Le test de langue première n'avait pas pour but de mesurer la compréhension de contenus propres à cette discipline, comme c'était le cas en mathématiques et en biologie. Les items devaient plutôt porter sur la compréhension de la langue définie selon la grille de compétences en langue première (cf. chapitre III) et qu'on se proposait d'évaluer à partir des supports de cours universitaires analysés dans le module A. Il était donc clair dès le départ que les compétences évaluées par le test de langue première allaient se différencier des savoirs et savoir-faire visés dans les épreuves d'examens de maturité bien davantage que ce n'était le cas pour les deux autres disciplines testées.

Pour effectuer le choix préalable du corpus d'étude, on a mis une priorité sur une documentation intervenant tôt dans les études, c'est-à-dire des supports de cours lus déjà au début du premier semestre des cours. A l'intérieur des ouvrages spécialisés retenus, on a mis l'accent dans toute la mesure du possible sur les premiers chapitres qui allaient être traités et travaillés au début du cycle des cours. Cette procédure de sélection, pour ainsi dire chronologique, permettait de cibler par les tests une matière aussi proche que possible des savoirs prérequis.

Pour élaborer des items dans le domaine «Vocabulaire», on a fait porter plusieurs questions sur des unités de signification codées au sein du savoir lexical spécifique lié au contexte des supports de cours à étudier. On a opté le plus souvent pour le format «choix multiple».

La construction des items portant sur la compréhension générale et la compréhension détaillée de l'écrit ainsi que sur les éléments de grammaire a commandé d'élargir l'attention au contexte global de l'unité de sens ou du groupement d'unités de sens codées. Afin de pouvoir créer les formats les plus appropriés, on a veillé, lors de la formulation de la question de l'item, à relâcher le lien étroit la rattachant à l'unité de sens codée. En d'autres termes, toutes les questions du test n'attendent pas nécessairement comme réponse correcte et explicite l'une des unités de sens codées en tant que telle, pas plus qu'elles ne s'y réfèrent obligatoirement; la solution peut se trouver aussi dans l'environnement proche d'une unité de sens codée.

Les lignes qui suivent présentent un exemple illustrant cette démarche qui va du support de cours à la question de détail posée dans un item.

La figure IV.4 ci-dessous est une reproduction d'un manuel utilisé comme support de cours en pédagogie: «*Lernen: Zwanzig Szenarien aus dem Alltag*» de Gerhard Steiner (2001); c'est un document de travail montrant le travail de codage du texte pour la sélection des questions en allemand. Les unités de sens codées ont été réunies dans un tableau (cf. figure IV.2 – en allemand – pour l'extrait du chapitre 1 du livre retenu). Les mots d'origine étrangère (Fw) du point de vue germanophone pouvaient être classés soit dans la discipline «Allemand», soit selon la langue d'origine respective; ils ont donc donné lieu à des codages mentionnant chacune de ces langues (français (F), italien (I), anglais (E), latin (L) ou grec (Gr)).

Figure IV.2: Extrait de la liste des unités de sens codées (mots d'origine étrangère en allemand - «Fremdwörter»)²⁴

F-I-E-L-Gr-Faktenwissen

AA:

2.5. Buch-/Skripttitel:

Steiner, G. (2001). *Lernen. Zwanzig Szenarien aus dem Alltag*. 3., version modifiée, Berne: Hans Huber.

Dateiname Kodierung: a1_paed_L_Steiner1_Kp1_Cl.rtf

F:	Komplexität	Experiment(2)	Gr:
Verhaltensrepertoires	Signal(tafel)(2)	experimentelle(r)(2)	lern theoretischen
E:	Organismus(5)	elaborierten	Theorie(n)(4)
distress	Lern prinzipien	existenzbedrohender	Grundmechanismus
stimulus	Neben prinzipien	Adaption	akustische
? behavioristische (n)(2)	Prinzip	Antizipation	? Lern theoretiker (4)
unconditioned stimulus	selektiv	verbal	? Kognition spsychologie
unconditioned response	präzise	Experimentator	Signal system(s) (4)
conditioned stimulus	? Kognitionspsychologie	Präsentation	psychischen
conditioned response	Inventar	Korrelaten	psychologischen
? Behaviorist	Konfiguration	verbal(er)(4)	Psychologie geschichte
L:	Signal system(s)(4)	Verbalisieren	drastisch
Labor(3)	antizipiert	Präzision	Phänomen
differenziert	? kognitiv(er)(3)	differenziertes	hirn physiologischen
Differenzierung	? kognitive(n/r) Prozesse(6)	Vokabular	Nervens system
visuell	klassische	Lokalisierung	physiologischen
komplexer(e/n)(4)		präzisen	

Comme on l'a indiqué plus haut avec la procédure chronologique, il s'agit d'un exemple pour le test concernant la première page du premier chapitre «Angst vor weissen Kitteln – Klassisches Konditionieren» (pages 15 à 27) à propos de laquelle il faisait sens de construire un item. Dans le court chapitre introductif 1.1 et dans le chapitre 1.2 qui traite de la découverte de Pavlov, des unités codées se trouvaient déjà, comme les signifiants «blouses blanches», «physiologie nerveuse», «théories» et «laboratoire» (pages 15), mais ces unités n'apparaissaient que trop rarement et dans un contexte peu favorable. Au chapitre 1.3 du livre, on trouve le paragraphe parlant des

²⁴ Le chiffre qui figure entre parenthèses (2) à côté d'une unité de sens codée indique le nombre d'occurrences, au sein du chapitre concerné, du terme en question, sans que celui-ci ait été accompagné d'une explication ou d'une définition. Par la mise en évidence en caractères **gras**, la personne chargée du codage signale quelle est la partie du mot (le lexème) qui se réfère à la langue d'origine concernée.

«réactions de peur comme éléments de comportement élémentaire» (cf. figure IV.4) qui comporte un premier groupement d'unités de sens assez dense (cf. soulignement et codes) comparativement à la longueur du texte.

Après avoir examiné si le contexte de ce groupement d'unités de sens donnait matière à des questions de tests, les collaborateurs ont déterminé, en une étape suivante, le registre de langue des unités de sens codées. Cette catégorisation s'est faite en allemand par le biais du dictionnaire électronique «Profile deutsch» (2002) qui est fondé sur le Cadre européen commun de référence CECR (2001), et qui permet d'accéder à environ 33 000 lemmes. Cet outil a permis de définir que l'unité de sens «differenziert» (Steiner, 2001, p. 16, alinéa 2, ligne 2) appartient au registre de langue C, mais que «Differenzierung» (ligne 5) se trouve déjà dans le registre D²⁵. On attribue aussi au niveau D l'unité de sens «(Verhaltens) repertoire(s)» (ligne 9). Les deux unités de sens ont été retenues chacune pour un item de test, (les items 1.1 et 1.2; cf. figures IV. 5 et IV.3).

Comme indiqué plus haut, la construction des items de vocabulaire ont ciblé l'attention sur l'unité de sens codée; à l'inverse, les items qui portaient sur la capacité générale de comprendre l'écrit ou la compréhension détaillée de l'écrit se fondaient sur un contexte plus large.

L'item 1.2 (figure IV.3) teste la maîtrise et l'emploi du vocabulaire en se fondant sur l'unité de sens «répertoire» déterminée à partir du texte de la figure IV.4. Il. La question porte sur la compréhension de ce mot et sur la capacité des bacheliers et des bachelières de le remplacer par un synonyme. Comme c'est souvent le cas pour des items de vocabulaire, l'item prend la forme d'une question à choix multiple.

Figure IV.3: Item de vocabulaire construit à partir du texte reproduit dans la figure IV.4 et des unités de sens recensées dans la figure IV.2.

Question 1.2

F_W_12

Quel substantif correspond le mieux au mot «répertoire» dans le contexte suivant?

Cochez la bonne réponse.

«Nous pouvons dire que cette réaction fait partie des potentialités comportementales originelles et naturelles; elle est une composante d'un répertoire dont les fondements sont innés.»

(Steiner, 2001, p. 16)

- 1 ☐ ...mécanisme
 - 2 ☐ ...ensemble
 - 3 ☐ ...instinct
 - 4 ☐ ...façon
-

²⁵ Le niveau D est un «titre de travail» choisi librement pour des termes qui n'ont pas d'entrée dans le grand dictionnaire de la langue allemande et dont on peut dès lors supposer qu'ils correspondent à un niveau d'exigence supérieur à C1 et à C2.

Figure IV.4: Reproduction du texte original (en allemand) codé par le concepteur du test avec les unités de sens codées, les registres de langue et le choix opérés pour les items du test (Steiner, 2001, p. 16–17)

1.3 Klassisches Konditionieren von Emotionen

Die Situation

Eine junge Mutter sitzt mit ihrem knapp anderthalbjährigen Kind im Wartezimmer des Augenarztes. Nachdem letztes Mal die versproften Tränenkanäle gespült werden mußten, steht heute lediglich eine Nachkontrolle an. Das Kind ist viel unruhiger als sonst, aber die Mutter lenkt es mit Geschichtenerzählen geschickt ab. Wie nun die Ärztgehilfin eintritt, beginnt das Kind wie am Messer zu schreien und zu strampeln. Das ist seine Reaktion auf die neue Reizsituation, die mit dem Eintreten der Ärztgehilfin, die das Kind zuvor nicht kannte, eingetreten ist. Die Reaktion ist völlig eindeutig; sie ist der Ausdruck für eine starke Emotion, für Angst, für Furcht vor etwas, vielleicht auch für einen Widerwillen gegen etwas. Eine charakteristische Reizsituation löst eine ebenso charakteristische Reaktion aus. Verfolgen wir die Lerngeschichte dieser Reaktion!

Angstreaktionen als elementare Verhaltensweisen

Angst gehört zu den grundlegenden, schon früh zu beobachtenden Emotionen. Sie differenziert sich im Alter zwischen 3 und 6 Monaten zusammen mit dem Zorn und dem Widerwillen aus einer gemeinsamen Frühform, nämlich einem Unbehagen oder einer Unruhe (engl. disstress) heraus, die ihrerseits eine erste Differenzierung eines allgemeinen Erregungszustandes darstellt (nach Allport, 1961 bzw. Bridges, 1932). Wir können sagen, daß diese Reaktion zu den ursprünglichen und natürlichen Verhaltensmöglichkeiten gehört; sie ist Bestandteil eines zumindest in seinem Grundbestand angeborenen Verhaltensrepertoires.

Daß es sich um eine höchst elementare, natürliche Verhaltensweise handelt, ist für die klassische Konditionierung von großer Bedeutung, denn diese erfolgt nur dort, wo bereits ausgebildete elementare Verhaltensformen (z. B. Reflexe wie der oben erwähnte Reflex der Speichelsabsonderung) vorhanden sind, die durch neue Reize ausgelöst werden können.

Zunächst wenden wir uns der einfachen Frage zu: Welcher Reiz hat eigentlich das Weinen und Schreien des Kindes ausgelöst? – Wir müssen bekennen: Ganz genau können wir das gar nicht sagen! Wir wissen nur, daß sich mit dem Erscheinen der Ärztgehilfin die gesamte Reizsituation verändert hat, und zwar in vielfältiger Weise. Die Person ist ein visuell wahrnehmbarer komplexer Reiz (engl. stimulus); eine menschliche Gestalt in weißem Kittel, mit einem Kopf von bestimmtem Aussehen, mit einem ganz bestimmten Körperbewegungs- und möglicherweise einem ganz bestimmten Annäherungsmuster. Sie ist aber auch, wenn sie spricht, eine Quelle für akustische Reize.

Auswahl von Stimuluselementen in komplexen Alltagssituationen

Die aktuelle Reizsituation in der Arztpraxis ist gegenüber derjenigen in Pawlows Labor eine viel komplexere! Durch welche Teile dieser Reizsituation nun eine Reaktion wie die Angstreaktion des Kindes ausgelöst wurde, ist für den Lehrer der Tat schwer zu sagen. Mit diesem Problem mußten sich die Lehrer der ersten Hälfte unseres Jahrhunderts schon auseinandersetzen. Sie hatten erkannt, daß in Alltagssituationen nur ausnahmsweise einfache Einzelreize auftraten, von denen man mit einiger Sicherheit annehmen konnte, daß sie eine bestimmte Reaktion auslösten. Der Lerntheoretiker Edwin R. Guthrie (1886–1959) war sich der Komplexität der Reizsituation im Alltag durchaus bewußt und hat von der Aktivität des Organismus bei der Stimulusauswahl gesprochen und formuliert, daß das, was der Organismus erfasse, zum Signal für das werde, was getan wird (Guthrie, 1959, 186). Das heißt nichts anderes, als daß nicht eine ganze Reizsituation eine Reaktion auslöst, sondern nur das, was aus ihr erfaßt wird; das aber wirkt dann als Signal und löst eine entsprechende Reaktion aus.

Schon 1931 hat ein anderer angloamerikanischer Lerntheoretiker, Edward L. Thorndike (1874–1949), Lerngesetze und -prinzipien formuliert, von denen eines das unterschiedliche Hervortreten von Stimuluselementen betraf. Eines der Nebenprinzipien befaßt sich mit der Prävalenz von Elementen (siehe Bower & Hilgard, 1981, 27). Dieses Prinzip besagt, daß ein Lerner selektiv auf besondere, d. h. herausragende oder auffallende Elemente der Reizsituation reagieren kann. Die auffallenden Teile der Reizsituation bieten sich dem Organismus gleichsam zum Erfassen an. Hier müssen wir einen bedeutenden Unterschied zwischen den Aussagen von Thorndike (1931) und Guthrie beachten: Während sich nach Thorndikes Ansicht die auffallenden Elemente einem passiven Organismus anbieten oder gar aufdrängen, spricht Guthrie ausdrücklich von der Aktivität des Organismus im Hinblick auf das, was von einer Reizsituation erfaßt werden soll. Guthrie weist also dem aktiven Organismus eine besondere Bedeutung gegenüber den Merkmalen der entsprechenden Reizsituation zu. Er sagt allerdings nichts Genaueres über diese Prozesse der Auslese von Reizmerkmalen; auch kommt er nicht darauf zu sprechen, daß allenfalls identische Teile von Reizsituationen für verschiedene Individuen völlig verschiedene Bedeutungen haben können und aus diesem Grund auch völlig unterschiedliche Reaktionen auslösen! Aufgrund der Pawlowschen Konditionierungstheorie oder der frühen behavioristischen Lerntheorien (wie denjenigen von Thorndike oder Guthrie) sind wir nicht in der Lage, präzise Vorhersagen über die Auswahl von Stimuluselementen (Teilen von Reizsituationen) zu machen; dazu sind Theorien notwendig, die etwas über die Aufmerksamkeitssteuerung zur Auswahl von Teilen von Reizen aussagen können, und das sind die neueren Theorien aus der Kognitionspsychologie (vgl. etwa Neisser 1976, 1979 oder Eysenck & Keane, 1990).

L'item 1.1 ci-après (cf. figure IV.5) centre précisément la question sur l'unité de sens codée (dans ce cas, les termes «différencier/différenciation») tout en se référant au contexte assez large, c'est-à-dire au paragraphe. Il n'est possible de comprendre les options proposées qu'après avoir saisi la logique de l'information de la phrase et l'avoir reliée à la question posée. Dans ce cas, une lecture détaillée de tout le contexte est une condition préalable à la découverte de la réponse correcte. Aussi cet item a-t-il été attribué à la compétence de compréhension détaillée de l'écrit. Une aide à la recherche réside dans le fait décrit que le désagrément est une forme primitive de l'émotion. L'erreur logique de la variante 2 doit donc faire écarter cette option.

Figure IV.5: Compréhension détaillée de l'écrit – Item construit à partir du texte reproduit dans la figure IV.4 et des unités de sens recensées dans la figure IV.2.

Question 1.1

Comment faut-il comprendre «différenciation» dans le contexte suivant?

Cochez la bonne réponse.

F_D_11

«La peur fait partie des émotions fondamentales observables dès le plus jeune âge. Entre 3 et 6 mois, elle se différencie de la colère et du dégoût à partir d'une forme primaire globale de détresse ressentie comme un malaise ou un déplaisir (angl. distress). Ce sentiment initial de déplaisir représente lui-même une première différenciation par rapport à un état général d'excitation (selon Allport, 1961, et Bridges, 1932).»

(Steiner, Gerhard. 2001. *Lernen: Zwanzig Szenarien aus dem Alltag*. Bern: Huber. Kap. 1: Angst vor weissen Kitteln – Klassisches Konditionieren, p. 16)

- 1 ☐ La peur se développe approximativement entre trois et six mois.
- 2 ☐ La peur se développe en même temps que la colère, le dégoût et le déplaisir.
- 3 ☐ La peur se développe en premier lieu comme un état primaire global.
- 4 ☐ La peur se développe à partir d'un état général d'excitation.

L'ensemble du test s'est articulé en dix parties distinctes correspondant aux différentes disciplines des universités suisses choisies dans le module A (cf. chapitre II). Chaque subdivision du test comporte au moins un texte de départ appartenant à la discipline d'études concernée. Les répondants doivent dans un premier temps lire le texte spécialisé; sur cette base, ils trouvent les items de test présentés en divers formats.

2.5.1.2 A propos de la traduction des items

S'agissant de la langue première, il convenait de tenir compte d'un facteur influant sur la comparabilité des valeurs testées: d'une part, les textes spécialisés pouvaient présenter un degré de difficulté variable selon les langues du fait de la nécessité où l'on s'était trouvé de traduire les items, d'autre part, les nuances de sens pertinentes pour tel ou tel item étaient parfois loin d'être déterminantes dans l'autre langue. Cette observation s'applique en particulier, en allemand, à certains mots d'origine étrangère qui sont pour beaucoup dans la difficulté des langages techniques spécialisés du point de vue germanophone, mais qui appartiennent plutôt à la langue standard en Suisse romande ou qui sont plus aisément identifiables du fait de leur étymologie plus familière.

A l'inverse, le fait que les textes de cours soumis aux étudiants romands et italophones étaient des extraits traduits, et non des textes nés dans la logique de leur langue a pu constituer un élément de difficulté supplémentaire pour les tests en français et en italien. Seule une analyse de texte plus approfondie permettrait d'évaluer plus précisément quelle influence exacte de tels facteurs ont pu avoir sur les valeurs du test. La recherche en lecture et en écriture connaît des méthodes, parfois assez anciennes, permettant de définir le degré de difficulté des textes (Fucks, 1955; Björnsson, 1968; Groeben, 1972; Bamberger & Vanecek, 1984; Tauber, 1984; Drinkmann & Groeben, 1989). Ainsi, l'enseignement de l'allemand et la discussion sur les standards valent actuellement une deuxième jeunesse à certains modèles (Köster, 2005, p. 37) qui dataient déjà: c'est le cas de l'index de lisibilité LIX, développé naguère dans les pays scandinaves par Björnsson (1968) et remis en selle par des logiciels informatiques permettant d'analyser les textes (par ex. le programme TextQuest, Social Science Consulting, 2007); dans les années 1980 en effet, cette manière venue du Nord d'évaluer la compréhensibilité d'un texte, fondée sur la «longueur moyenne des phrases» et la «proportion de mots longs et difficiles» avait eu du succès, vantée qu'elle était comme un moyen auxiliaire «simple, objectif et fiable» d'évaluer la lisibilité d'un texte (Commission pour les bibliothèques scolaires et les bibliothèques communales du canton de Berne, 1989, p. 3). Par contre, on ne sait pas dans quelle mesure ces analyses peuvent s'étendre à des types de textes universitaires, et en particulier à des approches plurilingues impliquant des traductions.

Il n'a pas été possible, par manque de ressources également du côté de nos partenaires, de suivre la procédure envisagée au départ, qui prévoyait de développer des items dans chacune des langues premières, puis de les faire traduire dans les deux autres langues. C'est donc uniquement l'équipe alémanique qui s'est consacrée à la construction des premiers items. Ce lot d'items en allemand a fait l'objet d'une première traduction, que les experts linguistes bilingues de nos partenaires romands et tessinois ont examinée et affinée. Pour des raisons linguistiques, il a fallu d'emblée mettre de côté certains items, en retravailler ou remanier d'autres. La sélection définitive a été opérée par les trois responsables principaux pour la langue première.

Les items prévus pour le test de langue première ont fait l'objet de prétests dans les trois régions linguistiques. Les résultats ont permis quelques comparaisons d'une langue à l'autre et, sur cette base, d'autres aménagements: dans une démarche commune de discussion et de recherche de consensus, on a optimisé encore des items ou des distracteurs que manifestement les Romands ou les Tessinois avaient parfois compris différemment que les Alémaniques dans la version originale. Cette procédure de mise à niveau de tous les items de langue a eu lieu en plusieurs étapes, autant lors de la construction du test que lors de la mise au point de la clé de correction.

Le test de langue première présente encore une autre particularité: il comporte une partie dévolue à la grammaire, donc aussi à des éléments d'orthographe et de ponctuation. Il n'était pas possible, pour des raisons de structure linguistique, de tester ce domaine technique à l'aide d'items traduits; il fallait au contraire travailler

exclusivement avec des items de tests construits à chaque fois dans la langue originale. Faute de moyens, nos partenaires romands et tessinois ont renoncé à élaborer des items originaux portant sur l'orthographe et d'autres compétences grammaticales (à cela s'ajoutait que les partenaires du Tessin avaient des objections de principe sur cette partie grammaticale). En fin de compte, il a été décidé, en considérant aussi les résultats de l'enquête du sous-projet A4, de maintenir malgré tout, pour la langue première «Allemand», une partie de test consacrée à la grammaire et à l'orthographe.

2.5.2 Description des cahiers de test

Pour la langue première «Allemand», on a produit cinq cahiers de test différents par le biais d'une matrice multicritères reliant les items entre eux dans une configuration tournante; pour le français et l'italien, ces cahiers étaient au nombre de quatre (sans items de grammaire). En d'autres termes, trois cahiers de tests étaient identiques dans les trois langues, seuls les cahiers quatre et cinq pour la langue première «Allemand» étaient différents puisqu'ils contenaient les items de grammaire qui venaient s'y ajouter.

Pour le test principal, 168 items en tout ont donc été construits. Les blocs des cahiers de test pour la langue première les critères ont été composés selon les critères de contenu et de forme suivants:

- Chaque cahier de test doit comporter des textes introductifs de diverses disciplines des études universitaires.
- La succession des items de lecture varie selon les cahiers, dans le cas idéal, un item donné figure autant de fois au début, au milieu et à la fin d'un cahier. Cette manière de procéder garantit que les sources de dérangement (par exemple la baisse de la concentration ou l'amélioration de la performance après la mise en route) ne se répercutent pas toujours sur les mêmes items.
- Les tests doivent être effectués en 22.5 minutes.

2.5.3 Caractérisation des items intégrés dans l'analyse générale et ancrage dans la grille de compétences

On a élaboré des formats fermés et ouverts pour tenir compte des différentes compétences partielles en langue première. Des 124 items de test intégrés dans l'analyse générale, 74 sont construits en format ouvert et 50 en format fermé à choix multiple. Les deux formats ouvert et fermé sont subdivisés selon les sous-formats ou les types de tâches. Parmi les 74 items de format ouvert, 81% sont ouverts et courts: il s'agit là d'items avec texte lacunaire où les bacheliers et bacheliers n'avaient qu'à inclure un mot manquant ou une terminaison. Pour les 14 items en sous-format «ouvert-long» (18.6%), il s'agit de questions plus exigeantes; il fallait, par exemple, attribuer des sous-titres à des paragraphes assez longs extraits de textes scientifiques, ou reformuler d'une manière grammaticalement correcte des phrases tirées d'un texte documentaire. Parmi les items à choix multiple, un peu plus des deux tiers se présentaient sous leur forme traditionnelle (cocher la variante correcte). On trouve en outre d'autres formats d'items comportant par exemple des tâches de classement (par ex. attribuer des titres données à des alinéas d'un texte).

L'analyse générale a donc intégré, pour la langue première, 77 items «nationaux», validés et statistiquement fiables autant pour l'allemand que pour le français. Pour la langue première «Allemand», on a ajouté au test 47 items tirés de la partie grammaticale, ce qui représente un total de 124 items de test. Leur répartition selon la grille de compétences est la suivante:

- Capacité générale de comprendre l'écrit (lire pour s'informer): 19 items (D_A_XX)
- Compréhension détaillée de l'écrit: comprendre des informations et des arguments: 40 items (D_D_XX)
- Vocabulaire: 19 items (D_W_XX)
- Compétences grammaticales: 17 items (D_G_XX)
- Orthographe: 29 items (D_G_XX)

La description des caractéristiques du test figure dans le rapport technique.

2.6 Le test de mathématiques

2.6.1 Précisions sur l'élaboration du test

2.6.1.1 A propos de la construction des items

L'exemple ci-après illustre la manière dont les items du test ont été construits à partir des unités de sens relevées dans les notes de cours des deux premiers semestres d'études universitaires.

La figure IV.6 reproduit un extrait d'une note de cours de l'EPFZ ayant servi de base à l'élaboration d'un item.

Figure IV.6: Extrait d'une note de cours originale de l'EPFZ («Physik, Felder, Geometrie», p. 30)

2.2 Raum 16.11.2004

Anwendung: Kartesische Gleichung einer Ebene. Ebene durch

$$P_0 = (x_0, y_0, z_0)$$

Normalvektor

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \neq \vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$



Ebene

$$\begin{aligned} &= \{P \in \mathbb{R}^3 \mid \overrightarrow{P_0 P} \perp \vec{a}\} = \left\{ (x, y, z) \mid \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 0 \right\} \\ &= \{(x, y, z) \mid (x - x_0)a_1 + (y - y_0)a_2 + (z - z_0)a_3 = 0\} \end{aligned}$$

Ebenengleichung:

$$xa_1 + ya_2 + za_3 = b$$

mit $b = x_0a_1 + y_0a_2 + z_0a_3$

Übung: Spiegelung s bezüglich der Ebene $\vec{a} \cdot \vec{x} = 0$ $\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

$$s(\vec{x}) = \vec{x} - \frac{2(\vec{a} \cdot \vec{x})}{\vec{a} \cdot \vec{a}} \cdot \vec{a}$$

2.2.2 Vektorprodukt

Definition: Vektorprodukt (Kreuzprodukt)

$$\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

$$(\vec{u}, \vec{v}) \mapsto \vec{u} \times \vec{v}$$

$$\vec{u} \times \vec{v} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ -u_1v_3 + u_3v_1 \\ u_1v_2 - u_2v_1 \end{pmatrix} = \begin{pmatrix} \det \begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix} \\ -\det \begin{pmatrix} u_1 & v_1 \\ u_3 & v_3 \end{pmatrix} \\ \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \end{pmatrix}$$

Les unités de sens principales (connaissances requises pour comprendre le contenu du cours) consistent ici dans la manière de décrire un plan et de le mettre en équation («équation cartésienne d'un plan»). Pour comprendre

l'équation d'un plan $xa_1 + ya_2 + za_3 = b$, on doit savoir que le vecteur $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ est perpendiculaire au plan défini. L'item construit sur la base de ces deux unités de sens est reproduit ci-dessous (figure IV.7).

Figure IV.7: Item correspondant à la note de cours de la figure IV.6

Le point $(x_0; y_0; z_0)$ et le vecteur $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ étant donnés, l'ensemble

$$\left\{ (x; y; z) \left| \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \circ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 0 \right. \right\}$$

représente un plan dans l'espace qui contient le point $(x_0; y_0; z_0)$.

Lequel des vecteurs ci-dessous est perpendiculaire à ce plan?

Cochez la bonne réponse:

1

$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$

2

$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$

3

$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$

4

$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$

M_G_11

2.6.1.2 A propos de la traduction des items

Les items ont été traduits en italien par une traductrice professionnelle. Son travail a été révisé par le concepteur des items, puis par la *Scuola Universitaria Professionale della Svizzera Italiana* (SUPSI). Le concepteur des items s'est lui-même chargé de les traduire une première fois en français. Son travail a été révisé par un spécialiste - francophone et bilingue - de didactique en mathématiques pour le niveau secondaire II. Celui-ci a aussi ajusté le contenu des items. Il s'est avéré que certaines notions mathématiques ne s'expriment pas de la même façon en Suisse romande qu'en France. Les termes peuvent même varier localement en Suisse romande. Le même phénomène a été constaté au Tessin. Des experts ont été appelés à la rescousse pour vérifier que les formulations étaient correctes dans plusieurs cas particuliers. Leurs commentaires et leurs corrections ont été intégrés au fur et à mesure dans la formulation des items. Dans certains cas, cela a même conduit à réajuster les items originaux en allemand. Les items traduits en français ont par ailleurs été testés auprès de deux classes à Neuchâtel. Les élèves n'ont pas fait de commentaires qui auraient rendu nécessaire une reformulation des items.

2.6.2 Description des cahiers de test

93 questions regroupant 175 items étaient à disposition pour élaborer la version définitive du test de mathématiques. Six cahiers formés de blocs de questions tournantes ont été constitués, en application de la configuration multi-matricielle qui a été choisie pour l'ensemble des tests. Les blocs de questions - et donc les cahiers du test - ont été formés sur la base des critères suivants:

141

- Les quatre types de questions (analyse, géométrie, stochastique, divers) devaient être représentés en quantités à peu près égales dans chaque bloc.
- Les trois formats d'items (voir point 2.6.3 ci-après) devaient être représentés en quantités à peu près égales dans chaque bloc.
- Le temps nécessaire pour traiter chaque bloc devait être à peu près équivalent.

2.6.3 Format des items et intégration dans la grille de compétences

La version définitive du test comprend des questions de trois formats différents:

(A) Questions à choix multiple (QCM) classiques

(B) Questions proposant plusieurs affirmations

(C) Questions ouvertes avec solutions univoques

Les questions sélectionnées pour le test se répartissent comme suit dans la grille de compétences:

- 29 items de calcul analytique (M_A_XX), 35 items de géométrie (M_G_XX), 24 items de stochastique (M_S_XX) et 29 items divers (M_X_XX).
- La répartition des items selon le type d'aptitudes testées est la suivante:

▪ Connaissances de définitions et de propositions	25 items
▪ Compréhension qualitative de notions mathématiques	49 items
▪ Calcul	27 items
▪ Maniement de termes et de formules mathématiques	20 items
▪ Interprétation de diagrammes	20 items
▪ Compréhension de textes	13 items
▪ Représentation visuelle	4 items

Certains items font appel à plusieurs types d'aptitudes. Les caractéristiques du test sont décrites dans le rapport technique.

2.7 Le test de biologie

2.7.1 Précisions sur l'élaboration du test

2.7.1.1 A propos de la construction des items

Tout le monde s'accorde à reconnaître que disposer de compétences en mathématiques renforce l'aptitude générale aux études supérieures. Les connaissances en biologie n'ont, elles, d'utilité vraiment évidente qu'en vue d'études dans le domaine des sciences expérimentales. C'est pour cette raison qu'une procédure spéciale a été mise en œuvre pour sélectionner les unités de sens à intégrer dans le test de biologie définitif (126 items). Trois critères de représentativité ont été appliqués successivement. Un premier tri s'est opéré en retenant les unités de sens qui présentent un caractère transversal (ensemble I). Le choix s'est ensuite porté sur les unités de sens qui surgissent particulièrement fréquemment dans les supports de cours (ensemble II). Finalement, les unités de sens particulièrement exigeantes ont été sélectionnées. Il s'agit d'unités de sens qui compliquent la compréhension d'un texte au point de représenter un obstacle à l'aptitude aux études supérieures (ensemble III). Cela dit, les unités de sens identifiées dans le module A ne couvrent pas complètement toutes les branches d'étude proposées dans l'ensemble des universités suisses (voir chapitre II). Le didacticien spécialisé chargé de sélectionner les unités de sens avait donc la possibilité d'en repêcher certaines, qui échappent aux trois critères de base, mais constituent des notions de biologie absolument centrales (des «classiques» de l'édifice) qu'il aurait été inconcevable de ne pas soumettre à la sagacité des élèves dans le cadre de cette étude.

Centré sur l'unité de sens «évolution», l'exercice suivant (figure IV.9) a par exemple été construit à partir d'une note de cours utilisée dans l'enseignement de la science des médias et de la communication.

L'exemple suivant, centré sur l'unité de sens «évolution», provient de l'ensemble de sources à caractère transversal (voir figure IV.8; les annotations manuscrites montrent le travail de codage). Il s'agit d'une page tirée d'une note de cours en sciences des médias et de la communication (en allemand) dans laquelle l'unité de sens «évolution» apparaît.

Figure IV.8: Page tirée d'une note de cours en sciences des médias et de la communication

Ein weiterer, mit dem Emergenzbegriff aufs Engste verbundener Begriff, ist der des sozialen Systems. Auch dieser Begriff ist bereits zu einem frühen Zeitpunkt innerhalb der gesellschaftstheoretischen Tradition aufgetaucht, nämlich bei Herbert Spencer.

Ein als "System" bezeichneter Realitätsausschnitt (es kann sich hierbei bspw. um eine Organisation oder eine ganze Gesellschaft handeln) wird als Untersuchungsgegenstand bestimmt, um anschließend dessen Elemente (soziale Handlungen, Kommunikation) samt ihren Eigenschaften und Beziehungen untereinander oder zum übergeordneten System zu untersuchen.

Soziologisches System nach Herbert Spencer

Exkurs "Herbert Spencer und das soziale System"

Herbert Spencer verwendet gleich wie Comte einen organischen Gesellschaftsbegriff. Darüber hinaus wird Gesellschaft bei Spencer, ebenfalls in Anlehnung an die Biologie, bereits explizit als "emergentes System" konzipiert. Das soziale System "Gesellschaft" gleicht dem biologischen System dadurch, dass es aus lebenden Teilen besteht. Die Teile, gemeint sind Menschen in ihrer Zugehörigkeit zu sozialen Schichten, Berufen, Altersklassen usw., sind einander ungleich und übernehmen unterschiedliche Funktionen. Durch die Erfüllung ihrer unterschiedlichen Funktionen ermöglichen die Teile das Fortbestehen des Ganzen und der übrigen Teile. Die Qualität und die Dauerhaftigkeit der arbeitsteiligen Beziehungen zwischen den Teilen macht die Stabilität der emergenten Ordnung des Ganzen aus. Stabilität ist bei Spencer ein zentrales Kennzeichen sowohl organischer wie auch sozialer Systeme.

Bei Spencer kommt noch ein weiteres Bestimmungsmoment sozialer Systeme hinzu - ein Moment, das bei Comte in dieser Weise noch nicht zu finden ist: gleich wie organische Systeme unterstehen soziale Systeme den Prinzipien der Evolution. Der Prozess der sozialen Evolution entspricht Spencer zufolge einem kontinuierlichen gesellschaftlichen Wachstumsprozess. Die evolutionäre Entwicklung in Richtung grösserer Komplexität untersteht dabei den gleichen Selektionskriterien wie die biologische Evolution nach Darwin. Herbert Spencer gilt denn auch neben John Stuart Mill als Begründer des ebenso bekannten wie problematischen Sozialdarwinismus (vgl. Spencer 1914). Der populäre Leitsatz "Survival of the Fittest" stammt im übrigen nicht von Charles Darwin, sondern von Herbert Spencer.

3!

5

L'unité de sens «évolution» apparaît quatre fois.

Lors de la construction de l'item, le concepteur du test devait tout d'abord estimer dans quelle mesure l'élève devait avoir compris et assimilé la notion d'évolution, dans le sens de la théorie de l'évolution, pour comprendre le passage sans recourir à un dictionnaire. Il a jugé qu'une compréhension approfondie n'était certainement pas nécessaire en l'occurrence pour comprendre le texte sans qu'il faille rappeler les deux grands fondements de la théorie de l'évolution ci-après, tels qu'ils sont enseignés en biologie (d'après Mayr, 2001):

- Une population ou une espèce se transforme du fait qu'elle produit en permanence de nouveaux génotypes et que la plupart des génotypes sont éliminés d'une génération à l'autre.
- Les génotypes disparaissent soit parce que le phénotype d'un génotype ne survit pas (sélection naturelle), soit parce que le phénotype d'un génotype ne parvient pas ou pas suffisamment à se reproduire (sélection sexuelle).

Dans l'exemple de texte reproduit ci-dessus, les associations (implicites) ci-après devraient amplement suffire:

1. Le processus de l'évolution est «aveugle» dans le sens qu'il se déroule sans atteindre un objectif ou sans «vouloir» atteindre un objectif (cf. proposition 1 de l'exercice), et sans non plus être «dirigé» par une force extérieure vers un but déterminé (cf. proposition 3).
2. La description du phénomène de l'évolution se distingue d'une théorie, qui tend à décrire le plus précisément possible les processus sous-jacents (cf. proposition 4).
3. Sans variation, il n'y a pas de sélection possible (cf. proposition 2).

A partir de ces prémisses, l'élève devrait être en mesure de conclure (implicitement) que l'évolution sociale

1. ne doit pas obligatoirement tendre vers une certaine forme (idéale) de société,
2. ne doit pas obligatoirement faire l'objet d'une théorie du développement social, même si on observe d'une certaine façon «quotidiennement» le phénomène du développement social dans le monde entier,
3. se fonde sur une grande variété de représentations de la vie en société (variation) qui se répandent (sélection positive), perdurent un certain temps et peuvent à nouveau disparaître avec le temps (sélection négative au sens d'une élimination).

Si ces trois associations sont activées, l'auteur du cours a réussi, grâce à sa référence à la notion biologique de l'évolution, ou plutôt à l'utilisation métaphorique de cette dernière, à exprimer en peu de mots des aspects essentiels de sa représentation de l'évolution sociale.

La formulation concrète de l'item tend dès lors à examiner si le bachelier ou la bachelière n'opère pas d'associations «malheureuses» à la lecture de la note de cours, c'est-à-dire des associations non souhaitées par l'auteur du texte didactique (voir figure IV.9):

Figure IV.9: Question du test relative à l'extrait de la note de cours reproduit dans la figure IV.8

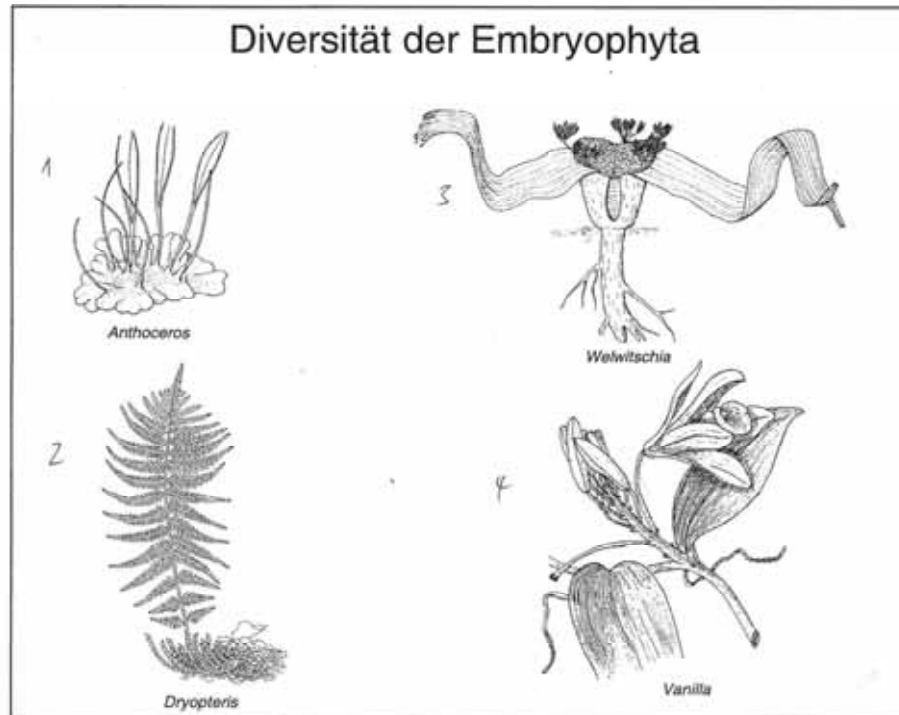
Test-Item B_EVO_01

Quelle est la condition préalable centrale pour tout type d'évolution?

- ☐ Il doit y avoir un but, vers lequel l'évolution doit se diriger.
- ☐ Il doit y avoir une diversité et un choix doit être possible.
- ☐ Il doit y avoir une force extérieure qui dirige les processus de l'évolution.
- ☐ Il doit y avoir une théorie indiquant comment les processus de l'évolution doivent se dérouler.

L'exemple suivant porte sur les unités de sens «spores» et «graines», de l'ensemble des sources «particulièrement exigeantes» des études de biologie. Tiré du livre de cours *Biodiversität: Pflanzen, Wirbeltiere, Pilze (Teil Pflanzen)*, il suppose que les étudiants aient appris avant l'université la distinction fondamentale entre deux notions et l'aient enregistrée dans leur mémoire à long terme: en effet, la nécessité de la parenthèse au septième paragraphe (voir figure IV.10) n'a de sens que si l'on connaît la différence entre graines et spores.

Ab. 4 Im Folgenden werden die vier Hauptgruppen der Landpflanzen (Bryophytina, Pteridophytina, Gymnospermae, Angiospermae) ausführlich vorgestellt. Dabei wird versucht, die Entstehung des komplexen Baus moderner Landpflanzen aufzuzeigen. Welches ist z.B. der Ursprung des Blattes oder der Blüte? Man nimmt an, dass die ersten Landpflanzen einfach organisiert waren, und es gibt verschiedene Theorien, wie im Verlauf der Evolution die Komplexität moderner Pflanzen entstehen konnte.



Ab. 5 Moose, Farne, Gymnospermen und Angiospermen unterscheiden sich aufgrund ihrer Lebensweisen, wobei teils der Gametophyt (Moose), teils der Sporophyt (Farne, Samenpflanzen) im Vordergrund des Lebenskreislaufes steht.

Ab. 6 1. Moose sind niedrig wachsende Pflanzen die häufig an schattigen, feuchten Standorten leben. Einige Arten sind befähigt, Trockenzeiten zu überdauern. Moose besitzen nur Rhizoide und einfache Leitstränge (keine Wurzeln und komplexen Leitbündel). Die Sporophyten bilden sich stets auf den grünen, dominierenden Gametophyten und bleiben mit diesem Zeitlebens verbunden. Es sind ca. 20'000 Arten bekannt, aufgeteilt in drei grosse stammesgeschichtlich getrennte Gruppen.

Streckbrief
Moose

Ab. 7 2. Bei den Farne dominiert der Sporophyt, der eine selbständige grüne Pflanze darstellt, die in Wurzel, Achse und Blätter gegliedert ist. Die Sprossachse hat meist kein sekundäres Dickenwachstum so dass die Pflanzen nicht beliebig hoch wachsen können. Farne leben meistens an relativ feuchten, schattigen Orten. Der Sporophyt bildet haploide Sporen (keine Samen), die die Ausbreitung ermöglichen. Es gibt keine Trennung von Genfluss und Ausbreitung. Die Gametophyten (Prothallien) erreichen höchstens einige cm Durchmesser. Die Gruppe umfasst ca. 9'000 Arten aufgeteilt in zwei Evolutionslinien (Clades) sowie mehrere ausgestorbene Gruppen (Paläophytikum, v.a. Devon).

Streckbrief
Farne

Ab. 8 3. Gymnospermen (= Nacktsamer) sind in Wurzel, Achse und Blätter gegliedert, und weisen sekundäres

Streckbrief
Nacktsamer

Les étudiants de premier semestre en biologie testés dans le module A doivent apporter ce savoir au risque de ne pas saisir un passage important du cours. Le texte souligne une différence importante avec un point d'exclamation, mais ne l'explique pas. La connaissance de la différence entre graines et spores constitue donc de toute évidence un savoir prérequis.

L'item reproduit dans la figure IV.11 vise à tester si la bachelière ou le bachelier serait en mesure, une fois à l'université, d'identifier ce qui distingue les graines des spores. En l'occurrence, c'est le nombre de cellules qui est déterminant. Les propositions de réponse 3 et 4 dans l'exercice sont donc fausses. Dans un deuxième temps, il s'agissait de tester si le bachelier ou la bachelière, dans le doute, saurait deviner que les spores sont des organismes plus élémentaires que les graines (elles sont unicellulaires). Une contrevérité dans la formulation de la proposition 1 les aide à trouver la solution. La proposition 1 ne peut pas être correcte car les organes sont constitués de tissus. Or ces derniers sont pluricellulaires par définition. Les organes le sont donc forcément aussi. Par élimination, la bonne réponse est donc la proposition 2.

Figure IV.11: Question du test relative à l'extrait de cours reproduit dans la figure IV.10

Test-Item B_GEN_16

Comment se distinguent les spores des graines?

- ☐ Les spores sont toujours pluricellulaires; les graines sont des organes unicellulaires des plantes à graines.
- ☐ Les spores sont toujours unicellulaires; les graines sont des organes pluricellulaires des plantes à graines.
- ☐ Les spores sont produites par les mousses, les fougères et les arbres; les graines sont produites par les plantes à graines.
- ☐ Les graines sont produites par toutes les plantes terrestres; les spores ne sont produites que par les algues.

2.7.1.2 A propos de la traduction des items

Tous les exercices ont été traduits deux fois en français. Une troisième personne, qui a étudié la biologie et dont la langue maternelle est également le français, a comparé les deux versions. Elle y a pioché les meilleurs passages pour mettre au point la version finale de l'exercice. Le processus de traduction en italien s'est déroulé de manière analogue.

Dans un second temps, les items traduits ont été soumis pour contrôle à trois élèves du Collège du Sud à Bulle. L'auteur des items en allemand et l'un des traducteurs ont pris note sur le vif des remarques des élèves sur la qualité linguistique des questions. Quelques subtilités linguistiques ont été intégrées ponctuellement dans les exercices sur cette base. Ils se sont assurés aussi que – question beaucoup plus importante – la bonne réponse dans les questions à choix multiple soit clairement identifiable pour les élèves dotés des connaissances nécessaires. Il s'est avéré que cela était déjà le cas, de façon optimale. Les traductions en italien n'ont pas été validées selon une procédure analogue par manque de temps et en raison d'objections méthodologiques élevées au Tessin. En outre, on est parti de l'hypothèse que, comme en français, les bonnes réponses étaient formulées d'emblée de sorte qu'elles soient clairement identifiables.

2.7.2 Description des cahiers de test

Sept cahiers formés de blocs de questions tournantes ont été constitués en biologie, en application de la configuration multi-matricielle qui a été choisie pour l'ensemble des tests (voir tableau IV.2 au point 2.1). Les blocs de questions – et donc les cahiers du test – ont été formés sur la base des deux critères suivants:

- Chaque bloc devait comprendre trois questions de chacun des sept domaines thématiques.
- Le temps nécessaire pour traiter chaque bloc devait être à peu près équivalent.

2.7.3 Format des items et intégration dans la grille de compétences

91 items ont été retenus dans la sélection large d'exercices en biologie. 85 d'entre eux revêtent la forme d'une question à choix multiple avec quatre réponses possibles, dont une seule est correcte. Les six autres items sont formulés de façon ouverte et appellent comme réponse un terme technique de biologie.

Les items se répartissent de la façon suivante sur les catégories de la grille de compétences (entre parenthèses: le système de numérotation des items dans le domaine concerné):

- Cytologie/Anatomie/Métabolisme: 31 items (B_FKT_XX, B_STW_XX);
- Traitement de l'information, comportement et immunobiologie: 17 items (B_INF_XX);
- Génétique/Biologie du développement: 18 items (B_GEN_XX);
- Ecologie: 15 items (B_OEK_XX);
- Systématique/Evolution: 10 items (B_EVO_XX).

La répartition des items par niveau d'exigence est le suivant:

- Niveau I: 47 items
- Niveau II: 44 items

Le test ne comprend aucun item du niveau d'exigence III. L'évaluation et le codage d'exercices où la créativité joue un rôle aussi important auraient été très gourmands en ressources et difficiles à objectiver. Cela aurait impliqué de demander aux élèves de produire un essai élaboré pour tester leur capacité de traiter des problèmes complexes de façon autonome dans le cadre d'un processus créatif et itératif à plusieurs niveaux (par exemple mettre au point une configuration expérimentale permettant de répondre à un certain type de problèmes ou se pencher de façon critique sur un phénomène en biologie).

Par rapport à la langue première et aux mathématiques, la biologie prépare moins bien les élèves à une formation académique dans n'importe quelle branche. Dans cette discipline, il est aussi moins commode de relier les supports de cours à l'origine du test aux branches universitaires examinées dans les sous-projets A1 et A3. 81 des 91 items testent des connaissances qui ne sont pré-requises qu'en biologie et en médecine à l'université (du moins parmi les branches analysées). Dans 38 cas, il s'agit de connaissances qui sont présumées acquises de manière *particulièrement fréquente* dans ces deux branches d'études. Les 43 autres items reposent sur des unités de sens spécialement exigeantes ou sur des connaissances sans lesquelles il est pratiquement impossible de comprendre les supports de cours utilisés à l'université dans ces deux branches.

Les caractéristiques du test sont décrites dans le rapport technique.

2.8 Le test des compétences transversales (TCT)

2.8.1 Description des bases du test

Le test d'aptitude utilisé en Suisse pour l'admission aux études de médecine se compose des sous-tests suivants:

1. Problèmes quantitatifs et formels
2. Travail avec soin et concentration
3. Notions élémentaires en médecine et en sciences expérimentales
4. Reconnaissance de fragments de figures
5. Diagrammes et tableaux
6. Compréhension de textes
7. Mémorisation de faits
8. Figures tubulaires
9. Mémorisation de figures

Le Centre pour le développement de tests et le diagnostic (CTD) est responsable du test d'aptitude aux études de médecine en Suisse. Il a construit pour l'étude EVAMAR II quatre blocs d'items à résoudre en 22 minutes et demie. Les items ont été choisis parmi les sous-tests 1, 3, 5 et 6 du test d'aptitude aux études de médecine, qui

sont les plus efficaces pour prédire le succès futur des études (Hänsen & Spicher, 2001). Le tri s'est fait en veillant à ce que les items soient suffisamment sélectifs et d'un degré de difficulté adéquat.

2.8.2 Description des cahiers de test

Quatre cahiers formés de blocs d'exercices tournants ont été constitués, en application de la configuration multi-matricielle qui a été choisie pour l'ensemble des tests. Chaque cahier de test du TCT comprenait deux blocs d'exercices choisis parmi les suivants (entre parenthèses: le numéro du bloc et le système de numérotation des items à l'intérieur du bloc):

Bloc 1: quantités et formes (T_QUANT), 10 exercices (TCT 1)

Bloc 2: compréhension de textes (T_TEXT), 12 exercices (TCT 2)

Bloc 3: notions élémentaires de sciences expérimentales (T_NAT), 10 exercices (TCT 3)

Bloc 4: diagrammes et tableaux (T_DIAGR), 10 exercices (TCT 4)

2.8.3 Format des items

Le test des compétences transversales a été examiné sous l'angle de sa conformité aux exigences statistiques minimales décrites au point 2.4, tout comme les tests des trois disciplines étudiées dans le cadre d'EVAMAR II.

Tous les exercices remplissent ces critères. Les items utilisés avaient du reste déjà fait leurs preuves dans d'autres contextes. La question d'éventuels problèmes de traduction ne se pose pas car seuls les élèves alémaniques ont effectué le TCT.

Tous les exercices du TCT sont des questions fermées. Les caractéristiques du test sont décrites dans le rapport technique.

2.9 Le questionnaire complémentaire

Les bacheliers et bacheliers ont rempli le questionnaire complémentaire à l'issue des tests. Il s'agissait donc de la dernière étape de la collecte de données organisée dans le cadre du module C. La version allemande contenait davantage de questions car les écoles alémaniques avaient davantage de temps à disposition. Le questionnaire a été conçu en allemand sur la base d'items collectés dans toutes les régions, pour être ensuite traduit en français et en italien. La fiabilité statistique des résultats a donc dû être examinée séparément pour chaque région, spécialement lorsque les réponses étaient saisies sur des échelles de mesure (voir le rapport technique sur ce point).

2.9.1 Raison d'être et contenu

Le questionnaire complémentaire avait pour but d'obtenir des informations sur l'échantillon qui aillent au-delà des données administratives fournies par les écoles. Il s'agissait en particulier de mieux cerner l'origine sociale de la famille des bacheliers et bacheliers, s'ils sont issus de l'immigration et quelles sont leurs habitudes linguistiques. Toutes ces données sont d'importantes variables de contrôle dans le cadre de l'évaluation des performances. Certaines questions particulièrement pertinentes du point de vue de l'aptitude aux études supérieures ont aussi été intégrées dans le questionnaire. Elles portaient notamment sur les compétences transversales, la personnalité et la motivation des élèves (voir point 1.2.1). Par manque de ressources, il fallu se contenter de données fournies directement par les élèves par écrit, plutôt que de les collecter selon des procédures plus objectives comme l'*assessment* ou l'observation. Les compétences qui avaient déjà été étudiées dans le cadre d'EVAMAR I (Ramseier et al., 2004) ou du rapport de Notter & Arnold sur la transition entre le gymnase et les études supérieures (2006) n'ont pas été intégrées dans le questionnaire complémentaire. L'accent a été mis sur des compétences qui semblaient bien se prêter à être mises en relation avec les

performances aux tests. Les bacheliers et bacheliers ont été interrogés aussi sur leur perception des tests et sur leurs projets d'avenir. En Suisse alémanique, les élèves ont par ailleurs été questionnés sur leur intérêt pour les cours donnés dans les trois disciplines évaluées dans le cadre d'EVAMAR II et sur le temps qu'ils consacrent à leurs devoirs à la maison dans ces mêmes disciplines (mathématiques, biologie, allemand). Quelques questions portaient aussi sur leur intérêt pour la lecture et sur leur perception de la valeur propédeutique du gymnase. Ces différents domaines sont décrits plus précisément ci-dessous.

2.9.2 Description du questionnaire

Plusieurs questions visaient à cerner le sérieux avec lequel les bacheliers et bacheliers ont participé aux tests. Ils devaient dire s'ils trouvaient les tests intéressants (de 1 = pas intéressant à 4 = très intéressant). Les efforts qu'ils disent avoir déployés pour remplir les tests de manière consciencieuse, engagée et concentrée ont été évalués sur une autre échelle de mesure comptant quatre niveaux.

Le volet «Ecole et enseignement» traite de l'articulation entre l'option spécifique et l'option complémentaire. Il englobe aussi une série d'items portant sur les compétences transversales. Les bacheliers et bacheliers devaient auto-évaluer leurs compétences et leurs stratégies d'apprentissage dans différents domaines: la planification de leur temps de travail, l'efficacité de leur préparation des examens, la gestion de leur temps, leur concentration, leur persévérance, leur esprit d'analyse et de synthèse, leur capacité d'argumenter et leur capacité d'adopter d'autres points de vue que le leur.

Dans la version alémanique du volet «Ecole et enseignement», des questions portent aussi sur le temps consacré à la maison aux devoirs de mathématiques, d'allemand et de biologie. Des items mesurent en outre l'intérêt et la motivation des élèves pour les trois branches testées. Une des questions était formulée ainsi: «Apprend-on tout en s'amusant aux cours de mathématiques ?» (réponses possibles de 1 = pas du tout à 4 = tout à fait). Les bacheliers devaient indiquer aussi si – et si oui à quelle fréquence – il leur arrive d'exercer à l'école certaines activités ayant une valeur propédeutique en vue d'études supérieures (comme trouver et traiter des informations utiles pour un travail ou préparer une expérience scientifique). Une autre série de questions les invitaient à évaluer dans quelle mesure ils sont armés pour relever le défi d'études supérieures. Ils devaient dire par exemple si le gymnase les avait préparés à aborder une question de façon scientifique ou à analyser un texte de façon systématique.

L'ensemble des bacheliers et bacheliers ont été interrogés sur leur intention de se lancer dans des études supérieures. Ils ont dû préciser dans quelle branche, auprès de quelle institution et s'ils projetaient d'étudier une année à l'étranger. Les élèves qui ne prévoyaient pas de poursuivre leurs études devaient indiquer leurs projets également. En Suisse alémanique, les bacheliers et bacheliers ont dû dire en complément s'ils envisageaient de participer au test d'aptitude aux études de médecine et s'ils avaient prévu de se préparer à ce test d'une façon particulière.

La dernière partie du questionnaire porte sur des données personnelles: sexe, lieu de naissance des parents, date d'immigration en Suisse le cas échéant, langue principale parlée à la maison et au sein du cercle d'amis. Pour pouvoir préciser leur origine sociale, les élèves ont été interrogés sur la formation de leurs géniteurs, sur le stock de livres disponibles au domicile parental et sur divers aspects de leur environnement quotidien qui dénotent un bon niveau socioculturel (une chambre à soi ou la présence d'œuvres d'art). En Suisse alémanique, des questions portaient aussi sur l'intérêt des bacheliers et bacheliers pour la lecture et sur le travail qu'ils accomplissent hors du cadre scolaire.

2.9.3 Qualité technique du questionnaire

Certains outils et certaines échelles utilisés dans le questionnaire complémentaire avaient déjà servi dans d'autres études, comme PISA (voir Kunter et al., 2002), BIJU (*Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter*) et TOSCA (pour BIJU et TOSCA: voir Trautwein et al., 2006). Certains items ont été développés spécialement pour EVAMAR II lors de la phase des pré-tests (voir point 2.2). Ils ont été

intégrés ensuite dans le test sous des formes adaptées. A quelques rares exceptions près, les échelles affichent des valeurs d'homogénéité satisfaisantes (coefficient alpha de Cronbach). En général, elles atteignent les mêmes ordres de grandeur que les échelles d'origine. Lorsque l'on prend les régions linguistiques isolément, les valeurs d'homogénéité se ressemblent fortement aussi, quel que soit l'échantillon et l'échelle de mesure considérés. Ce résultat réjouissant n'allait pas de soi, car les items traduits n'avaient jamais été utilisés avant.

3 Définition de l'échantillon

3.1 Définition de la population cible

La population cible se compose des bacheliers et bachelères qui ont passé leurs examens de maturité en été 2007 en Suisse, à l'exclusion de ceux de Bâle-Campagne et du canton de Genève. Intégrer Bâle-Campagne dans la population cible aurait risqué de fausser les comparaisons car les examens ont lieu en automne dans ce demi-canton. A Genève, les lycées ont très vite élevé des objections massives contre l'étude. La plupart des établissements ont refusé de transmettre leurs données dans le délai qui aurait permis d'intégrer Genève dans la population cible (fin 2006). Pour diverses raisons, le groupe de conduite du projet EVAMAR II a renoncé à intervenir auprès des autorités compétentes pour forcer les gymnases à jouer le jeu. La validité de résultats obtenus sous la contrainte aurait été sujette à caution (risque de falsification important).

L'échantillon a été stratifié en sept blocs afin de reproduire aussi fidèlement que possible la structure très hétérogène du paysage gymnasial suisse mais aussi pour pouvoir procéder à des comparaisons fiables des sous-populations. Le premier bloc est composé du plus grand canton de Suisse de Zurich. Vu la taille du canton, les résultats qui y ont été obtenus se prêtent bien à des analyses structurelles internes, sans qu'il soit nécessaire pour cela de passer par une collecte de données auprès de l'ensemble de la population cible²⁶. La partie alémanique du canton de Berne constitue un bloc en soi. Elle est la seule zone en Suisse alémanique où la première des quatre années d'études conduisant à la maturité peut être accomplie dans une école du niveau secondaire I plutôt qu'au gymnase. Le bloc «Suisse alémanique, petit» a été formé avec les petits cantons alémaniques qui comptent moins de 15 classes de maturité (AI, NW, OW, GL, UR, AR, SH, VSd, SG). La création de cette strate devrait corriger le fait que les classes de ces petits cantons avaient une probabilité moindre d'être sélectionnés pour participer à EVAMAR II. Le bloc «Suisse alémanique, grand» est composé des cantons alémaniques résiduels (SZ, SO, TG, GR, BS, AG, SG, LU). Le bloc «Suisse romande 1» comprend les cantons romands où la durée réglementaire du gymnase à proprement parler se limite à trois ans (BEf, JU, NE, VD) et le bloc «Suisse romande 2» ceux où cette durée est de quatre ans (FR, VSf). Il était prévu également de constituer une strate avec le canton du Tessin.

Le tableau IV.3 montre comment les classes et les élèves de la population cible sont répartis par strates.

²⁶ Ces analyses sont réalisées à titre complémentaire. Elles ne sont pas prises en charge financièrement par le mandataire CDIP/SER.

Tableau IV.3: Population cible

Strates	Zones	Nombre de classes	Nombre d'élèves
1	Suisse alémanique 1: Zurich	114	2144
2	Suisse alémanique 2: BEd (trois ans de gymnase au minimum)	98	1927
3	Suisse alémanique 3: petits cantons (nombre de classes ≤ 15): AI, NW, OW, GL, UR, AR, SH, VSd, ZG	52	964
4	Suisse alémanique 4: grands cantons (nombre de classes > 15): SZ, SO, TG, GR, BS, AG, SG, LU	237	4595
5	Suisse romande 1 (trois ans de gymnase au minimum): BEf, JU, NE, VD	152	3133
6	Suisse romande 2: FR, VSf	53	1055
7	Tessin	41	775
Total		747	14 593

3.2 Constitution de l'échantillon

La sélection des bacheliers et bacheliers a eu lieu selon une procédure à un niveau. Les classes ont été choisies en proportion de leur taille (Cochran, 1977, p. 150; Kish, 1965, p. 182; Lehtonen & Pahkinen, 1995, p. 7 [*single-stage cluster sampling with probability proportional to size*]). Tous les élèves d'une classe sélectionnée devaient en principe participer aux tests. Le tableau IV.4 montre comment les classes et les élèves compris dans l'échantillon sont répartis par strates.

Tableau IV.4: Echantillon

Strates	Zones	Nombre de classes	Nombre d'élèves
1	Suisse alémanique 1: Zurich	75	1439
2	Suisse alémanique 2: BEd (trois ans de gymnase au minimum)	30	587
3	Suisse alémanique 3: petits cantons (nombre de classes ≤ 15): AI, NW, OW, GL, UR, AR, SH, VSd, ZG	30	578
4	Suisse alémanique 4: grands cantons (nombre de classes > 15): SZ, SO, TG, GR, BS, AG, SG, LU	30	590
5	Suisse romande 1 (trois ans de gymnase au minimum): BEf, JU, NE, VD	30	636
6	Suisse romande 2: FR, VSf	30	612
7	Tessin	35	654
Total		260	5096

3.3 Taux de retour

Le tableau IV.5 montre le nombre de classes et bacheliers et bacheliers qui ont participé à l'étude dans les différentes strates. Tous les élèves qui ont participé à au moins un des quatre tests sont comptabilisés.

Tableau IV.5: Niveau de participation des classes et des élèves selon la strate géographique

Strates	Classes par échantillon	Participation	Taux de retour des classes	Elèves par échantillon	Participation	Taux de retour des élèves
1 Zurich	75	67	89%	1439	1204	84%
2 Suisse alémanique, gym3	30	27	90%	587	496	84%
3 Suisse alémanique, petits	30	25	83%	578	459	79%
4 Suisse alémanique, grands	30	29	97%	590	560	95%
5 Suisse romande 1 (gym3)	30	27	90%	636	509	80%
6 Suisse romande 2	30	30	100%	612	545	89%
7 Tessin	35	10	29%	654	131	20%
Total	260	215	90%	5096	3904	77%

La collecte de données n'a pas pu avoir lieu comme prévu dans toutes les classes souhaitées. Cela ne pose toutefois un problème que dans le cas du Tessin. Les résultats de ce canton n'ont pas pu être pris en compte en

raison du faible taux de participation des écoles²⁷ et parce que, dans un cas, une école a choisi elle-même les classes qui devait participer au projet. Une sélection aléatoire n'était ainsi plus garantie. Le taux de retour se monte à 91% pour les classes et à 85% en termes d'élèves (sans le Tessin). Cette différence s'explique par les absences de certains élèves le jour des tests. Au total, 3800 personnes ont fourni des données exploitables.

La proportion de filles dans l'échantillon sélectionné pour les tests se monte à 58,2%, celle des garçons à 41,8%. Les taux étaient comparables dans l'enquête menée dans le cadre d'EVAMAR I. Ils sont très représentatifs de la population suisse dans son ensemble (voir Ramseier et al., 2004). D'autres spécificités structurelles de l'échantillon seront exposées plus loin dans le cadre des comparaisons des résultats par strates.

L'échantillonnage a lieu sur la base de groupes de tailles différentes. En l'espèce, la sélection s'est faite proportionnellement à la taille des classes. Les classes et les élèves n'avaient donc pas tous les mêmes chances d'être choisis pour participer aux tests comme cela aurait été le cas par exemple dans le cadre d'un processus de sélection aléatoire simple. Plus la classe est grande, plus sa probabilité est grande d'entrer dans l'échantillon. Dans ce genre de situations, les probabilités différentes d'être sélectionné doivent être prise en compte dans l'estimation des paramètres. C'est pourquoi les données ont dû être pondérées (voir Satin & Shastry, 1982).

4 Collecte et traitement des données

4.1 Déroulement de la collecte des données

Les élèves ont été soumis aux tests durant les trois dernières semaines de cours de leur dernière année de gymnase. Les calendriers scolaires n'étant pas harmonisés, la collecte des données s'est déroulée entre fin avril et début juillet 2007. La durée de la collecte a été comparable dans toutes les classes. Comme dans le cas d'HarmoS, les écoles se sont elles-mêmes chargées du déroulement des tests sur la base d'instructions précises. Les écoles tessinoises n'ont pas été en mesure de faire participer un nombre suffisant de bacheliers et bacheliers aux tests (voir point 3.3). Tous les documents nécessaires à la réalisation des tests, adressés nominativement, étaient prêts à être envoyés au moment où la décision de faire marche arrière a été prise.

4.2 Codage des réponses aux tests

Les réponses aux questions fermées ont été enregistrées telles quelles dans les bases de données électroniques. En revanche, il a fallu coder préalablement les réponses aux questions ouvertes. En biologie, cela ne concernait qu'un petit nombre d'exercices. En mathématiques, les réponses aux questions ouvertes pouvaient être codées objectivement, car les solutions correctes étaient chaque fois clairement définies. Les réponses au questionnaire complémentaire ont été saisies telles quelles, sans codage supplémentaire.

Les critères de codage sont moins évidents en langue première qu'en biologie ou en mathématiques. Les personnes chargées de corriger ce type d'exercices ont suivi une formation pointue. Leur travail a été soumis à des contrôles de qualité. Tous les exercices ont été corrigés par une deuxième personne. Les deux versions ont ensuite été confrontées. Les cas litigieux ont été tranchés en commun. La conceptrice du test a vérifié des échantillons prélevés tout au long des travaux de correction. Un examen de la fiabilité transversale du codage a en outre été fait. L'échantillon a été constitué de façon transversale, en tenant compte de la distribution des écoles dans les cantons et de la fréquence approximative des corrections par correcteur. L'examen a porté sur des travaux de correction de deux types d'exercices: ceux qui sont structurés de telle façon qu'ils doivent être

²⁷ Les effectifs de la population cible sont faibles au Tessin. De facto, il aurait fallu procéder à une enquête auprès de pratiquement toutes les classes pour obtenir de bonnes données. Cela posait problème en particulier du fait que les tests devaient se dérouler juste avant les examens de maturité. Cela aurait entraîné une charge beaucoup plus lourde pour le système gymnasial tessinois que dans les strates où seuls un nombre limité de classes ont participé aux tests.

corrigés par étapes (exercices de type *Partial Credit*) et ceux pour lesquels plus d'une seule réponse peut être considérée comme correcte et qui entraînent donc des travaux de correction complexes. L'examen de fiabilité transversale a été pratiqué à l'aide du coefficient IRC (*Item Response Curve*) de Fröh (2004, p.179). Les résultats sont bons voire excellents. Un seul item n'a obtenu une valeur «que» suffisante. Aucun item supplémentaire n'a donc dû être écarté du test sur cette base.

5 Résultats nationaux

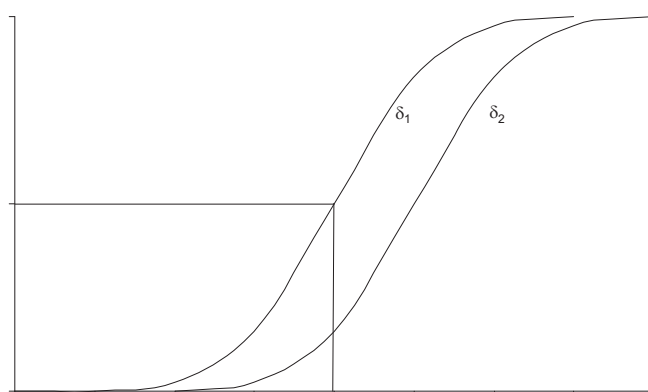
5.1 Généralités

Cette section expose les résultats de l'ensemble de l'échantillon. Deux méthodes de calcul ont été appliquées aux résultats des tests:

Méthode de calcul A: elle permet de mettre en évidence le taux global de réponses correctes dans un domaine thématique ou de compétence donné. Les résultats de l'ensemble des items sélectionnés pour un test ou sous-test sont simplement additionnés sur une base standardisée. Le maximum possible pour chaque cahier de test a été fixé à 20 points (10 points dans le cas de la biologie). Cette standardisation était nécessaire car les élèves ne devaient traiter que deux blocs d'exercices dans chacun des tests (et non tous) et que ces blocs ne comprenaient pas tous le même nombre d'exercices. Les blocs ne présentaient pas tous le même degré de difficulté, mais cela ne pose pas de problème à ce stade car les cahiers ont été répartis aléatoirement entre les élèves.

Méthode de calcul P: elle permet de se faire une idée encore plus précise des aptitudes personnelles des élèves en se fondant sur la théorie des probabilités. Les modèles probabilistes reposent sur l'hypothèse qu'il existe un lien stochastique entre le comportement de réponse d'une personne, ses aptitudes et la difficulté de l'exercice. Cette relation est exprimée graphiquement par une fonction logique mathématique non-linéaire connue sous le nom d'*Item Response Curve (IRC)* ou d'*Item Characteristic Curve (ICC)* (voir Rost, 2004, p. 96).

Figure IV.12: Item Characteristic Curve



La figure IV.12 illustre l'ICC de deux exercices présentant des degrés de difficulté différents. Elle montre la probabilité p qu'une personne θ résolve des exercices de difficulté δ_1 et δ_2 étant donné ses capacités (Kubinger, 2003, p. 3). Le modèle de Rasch, qui est appliqué ici, fait dépendre de deux paramètres (les aptitudes de la personne et la difficulté de l'exercice) la probabilité qu'une solution soit correcte. Plus une personne a de

grandes capacités, plus il est probable qu'elle résolve correctement l'exercice. Et plus un exercice est difficile, plus il devient improbable qu'une personne dotée d'aptitudes données en vienne à bout. Un grand avantage de ce modèle est que le résultat du test et le degré de difficulté de l'exercice peuvent être reproduits sur la même échelle. L'échelle est construite de manière à ce que la probabilité de trouver la solution correcte soit de 50% ($p = 0.5$) lorsque le degré de difficulté de l'exercice et les capacités de la personne sont en adéquation. La relation entre la probabilité de trouver la solution correcte, les aptitudes personnelles et la difficulté de l'exercice est d'une importance cruciale pour interpréter les résultats du test. Il est possible de déterminer pour chaque exercice la probabilité qu'un élève le résolve sur la base de son résultat individuel au test.

Une valeur de 0 calculée avec la méthode P correspond à des aptitudes personnelles d'un niveau moyen. L'échelle est ouverte vers le bas et vers le haut, mais les valeurs se situent en général entre -3 et +3. L'estimation des aptitudes personnelles s'appuie sur les exercices contenus dans tous les cahiers de test, en tenant compte de leurs degrés de difficulté différents.

Dans les diverses comparaisons entre sous-populations auxquelles il sera procédé au point 6 (ci-après), les aptitudes personnelles estimées selon le modèle de Rasch seront standardisées (moyenne = 500 points).

L'échantillon a été constitué à partir de classes et non directement à partir de bacheliers et de bacheliers (échantillon d'agrégats). Tous les élèves d'une classe choisie ont été inclus dans l'échantillon. La probabilité d'être sélectionné individuellement dépend donc de la taille de la classe. Plus elle est grande, plus la probabilité croît d'entrer dans l'échantillon. Les données doivent donc être pondérées. Le poids d'un échantillon correspond à l'inverse de la probabilité d'être sélectionné, après correction du taux de réponse (Satin & Shastry 1982, p. 35). Les cas où la probabilité d'être choisi est moindre sont donc surpondérés. C'est pourquoi le nombre de cas indiqué est systématiquement supérieur à la taille réelle de l'échantillon.

Comme cela a déjà été mentionné au point 3.3, les résultats du Tessin n'ont malheureusement pas pu être pris en compte.

5.2 Résultats de langue première

5.2.1 Résultats d'ensemble

Le tableau IV.6 montre les résultats d'ensemble obtenus au test de langue première. Pour des raisons techniques, le volet «Compétence grammaticale et orthographe» n'a été réalisé qu'en Suisse alémanique. Ses résultats sont indiqués séparément mais ils n'ont pas intégrés dans le total. Grâce à l'organisation réticulaire des cahiers de test, il est toutefois possible d'estimer aussi les aptitudes personnelles des élèves en Suisse romande en la matière (méthode de calcul P).

Tableau IV.6: Résultats du test de langue première

Types de résultats	Moyenne A	Ecart-type A	Maximum A	Moyenne P	Ecart-type P	Minimum P	Maximum P
TOT_LANGUE	31,22	9,27	60,00	,20	,562	-1,84	2,03
C-E_GEN	9,89	4,94	20,00	,38	,621	-2,11	2,30
C-E_DETAIL	10,52	3,97	20,00	,44	,676	-2,29	2,31
VOC	10,81	4,39	20,00	,39	,424	-1,18	1,90
GRAM	13,49	2,60	18,92	,38	,399	-1,36	2,32

N = 13 361 (pondéré)

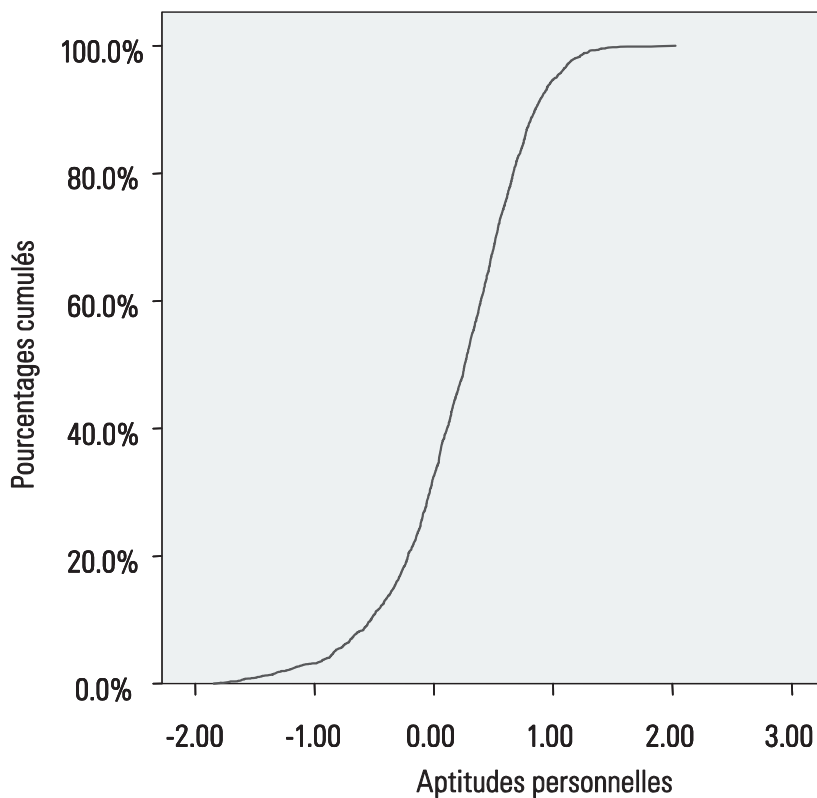
Explication des abréviations apparaissant dans le tableau:

- TOT_LANGUE: résultat global en langue première (hors grammaire)
- C-E_GEN: compréhension générale de l'écrit (lire pour s'informer)
- C-E_DETAIL: compréhension détaillée de l'écrit (information et argumentation)
- VOC: fonctionnement de la langue/vocabulaire
- GRAM: fonctionnement de la langue/compétence grammaticale et orthographe
- Moyenne A: moyenne des points obtenus selon la méthode de calcul A (maximum = 20 points par domaine, soit 60 points en tout)
- Ecart-type A: écart-type autour de la moyenne A
- Maximum A: nombre maximum de points atteint selon la méthode de calcul A
- Moyenne P: niveau moyen des aptitudes personnelles selon la méthode de calcul P (la valeur 0 correspond à des aptitudes « moyennes »)
- Ecart-type P: écart-type autour de la moyenne P
- Minimum P: niveau d'aptitudes personnelles le plus bas selon la méthode de calcul P
- Maximum P: niveau d'aptitudes personnelles le plus élevé selon la méthode de calcul P

Environ la moitié des exercices ont été résolus correctement en moyenne (beaucoup plus dans le domaine «Compétence grammaticale et orthographe»). En moyenne, les aptitudes personnelles estimées des bacheliers et des bacheliers sont telles qu'elles leur permettent de résoudre des items d'un niveau d'exigence supérieur à la moyenne dans tous les domaines de compétence. La dispersion des résultats des élèves est considérable. Au moins un élève a obtenu le maximum de 20 points dans tous les domaines, excepté dans le volet «Compétence grammaticale et orthographe». Globalement, la dispersion des résultats des classes est importante aussi (minimum: TOT_LANGUE = 20,72, GRAM = 7,55; maximum: TOT_LANGUE = 41,41, GRAM = 17,23). La meilleure classe a résolu correctement deux fois plus d'exercices que la plus mauvaise en moyenne.

La figure IV.3 représente graphiquement la distribution cumulée des aptitudes personnelles à la langue première, estimées globalement.

Figure IV.13: Estimation globale des aptitudes personnelles en langue première



5.2.2 Résultats d'items particuliers

Cette section s'intéresse à la manière dont une série d'items ont été résolus. Elle met en exergue la relation entre le taux de bonnes réponses et le degré de difficulté objectif de l'exercice (δ), estimé à l'aide du modèle de Rasch. Une personne qui disposerait d'aptitudes personnelles égales à celles calculées avec la méthode de calcul P aurait 50% de chances de résoudre l'exercice concerné (voir point 5.1). Une valeur de 0 pour δ correspond à un exercice de difficulté moyenne. L'échelle est ouverte vers le bas et vers le haut, mais les valeurs se situent en général entre -3 et +3,

Le premier exemple concerne un exercice de vocabulaire (voir figure IV.14). Il a été résolu correctement par 77,3% des bachelières et bacheliers ($\delta = -0,98$).

Figure IV.14: Exercice facile de vocabulaire (exemple)

Question 4.7

Quel adjectif peut remplacer «délibérées» dans la phrase suivante?

Cochez la bonne réponse.

F_W_47

«Les répercussions sur la façon de voir, les émotions et le comportement, qu'elles soient individuelles ou sociales, délibérées ou fortuites, à court ou à long terme, désirables du point de vue social, ou nuisibles ...»

(Bonfadelli, 2005, p. 12)

- 1 ☐ obtenues
- 2 ☐ produites
- 3 ☐ impliquées
- 4 ☐ voulues

Les élèves n'ont pas eu de peine à éliminer la proposition 1 «obtenues» (3,1% des réponses) et la proposition 3 «impliquées» (3,8% des réponses). La proposition 2 «produites» était un piège. Elle a été choisie relativement souvent (12,6% des réponses). Elle peut en effet sembler logique en tant qu'antonyme de «fortuites».

L'exercice présenté dans la figure IV.15 teste la compréhension détaillée d'informations et d'arguments écrits. Les élèves avaient à leur disposition un bref extrait d'un ouvrage sur l'histoire de la persécution des sorcières en Europe²⁸. À la lumière de ces informations, ils devaient rectifier une affirmation fautive. Les performances des bacheliers et des bacheliers montrent que l'exercice peut être considéré comme très facile ($\delta = -1,96$), ce qui n'était pas l'avis des experts. 89,8% des élèves ont trouvé la bonne réponse.

Figure IV.15: Exercice très facile de compréhension détaillée de l'écrit (exemple)

Question 5.2

L'affirmation suivante contient deux erreurs. Rectifiez-la en vous référant au texte 5.1.

Cochez les bonnes réponses.

Contrairement aux tribunaux civils, les tribunaux ecclésiastiques condamnaient la plupart des accusées à être brûlées sur un bûcher. F_D_52a

- 1 ☐ Les tribunaux ecclésiastiques condamnaient les accusées le plus souvent «à la preuve par l'eau», c'est-à-dire la mort par noyade.
- 2 ☐ Les tribunaux civils ne condamnaient aucune femme à mort.
- 3 ☐ Les tribunaux ecclésiastiques aussi bien que les tribunaux civils condamnaient des accusés à mort sur le bûcher.
- 4 ☐ Tout comme les tribunaux civils, les tribunaux ecclésiastiques condamnaient les accusées à mourir de faim.

²⁸ Note du traducteur: on trouve l'extrait en question (texte 5.1) dans la version allemande du présent rapport (Abb. IV.15, p. 148).

Le volet consacré aux exercices de grammaire n'a été soumis qu'aux élèves alémaniques. Il s'est avéré comporter beaucoup d'items très faciles à résoudre²⁹.

L'exemple suivant (figure IV.17) porte sur la compréhension générale de l'écrit. Il s'agit typiquement d'un exercice qui a causé beaucoup voire énormément de problèmes à de nombreux gymnasiennes et gymnasiens. La base de l'exercice est un petit texte relativement dense d'introduction à la science des médias³⁰. Les élèves étaient priés d'indiquer dans quel ordre un certain nombre d'éléments apparaissent dans ce texte. Pour réussir cet exercice, il fallait être capable de cerner le fil du récit dans le texte présenté, d'identifier les thèmes principaux de chaque paragraphe et de laisser tomber les détails insignifiants.

Figure IV.17: Exercice difficile de compréhension générale de l'écrit (exemple)

Question 4.3

Les indications suivantes figurent dans la marge du texte 4.1 dans le manuel.

Dans quel ordre apparaissent-elles?

- 1 association faïtière
- 2 non homogène
- 3 multiplicité des dénominations
- 4 différentes manières de voir
- 5 naissance de la branche

L'ordre correct des indications est le suivant:

(Inscrivez les numéros dans les cases)

☐ ☐ ☐ ☐ ☐

F_A_43

Cet exercice a ceci de spécifique qu'il force à structurer sa pensée de façon linéaire tout au long du texte. Il teste donc une compétence de base: l'aptitude à une lecture précise et non sélective (Kintsch, 1998; Gernsbacher, 1997). La conceptrice du test était d'avis qu'il s'agissait d'un exercice de difficulté moyenne voire plutôt facile car les contenus principaux du texte étaient indiqués. Les élèves n'avaient donc pas besoin de les identifier eux-mêmes. Étonnamment, seuls 21,9% des bacheliers et bacheliers ont résolu le problème correctement ($\delta = 1,77$). Une analyse détaillée de la fréquence des différents types d'erreurs pourraient fournir des informations intéressantes.

Les deux exercices suivants présentent un degré de difficulté important voire très important (figures IV.18 et IV.19). Ils testent également la compréhension générale de l'écrit. Mais les questions sont cette fois formulées de façon ouverte et les réponses sont évaluées de façon graduée sur une échelle de points (procédure de codage par «crédits partiels»).

Le premier exercice (figure IV.18) consistait à trouver un titre adéquat à un texte d'introduction à l'histoire de la médecine qui en survole les principales étapes, d'Hippocrate à la recherche sur le diabète au 20^e siècle. Le texte est de longueur moyenne et il est relativement simple à comprendre. L'exercice s'est avéré être l'un des trois plus difficiles parmi les questions à formulation ouverte finalement retenues dans l'enquête. Des points ont été attribués en utilisant une échelle à cinq niveaux: 1 = à moitié correct³¹, 2 = simple mais correct, 3 = correct et

²⁹ Note du traducteur: un tel exercice est présenté dans la version allemande du présent rapport (Abb. IV.16, p. 149).

³⁰ Note du traducteur: on trouve le texte en question dans la version allemande du présent rapport (Abb. IV.17, p. 150).

³¹ Ont été considérées comme «à moitié correctes», les réponses qui faisaient l'impasse sur des aspects essentiels du texte, qui mêlaient des éléments corrects et erronés ou qui, sans être fausses, ne restituaient pas la spécificité du texte (par exemple: «L'histoire des maladies humaines», «Les erreurs de la médecine»).

élaboré, 4 = correct et en faisant valoir des connaissances supplémentaires, 5 = correct et créatif sur le plan linguistique. L'appréciation la plus fréquemment délivrée est «à moitié correct» (39,2%), suivie de «simple mais correct» (18,8%). Les trois autres niveaux d'appréciation ont rarement été utilisées (niveau 3: 2% des élèves, niveau 4: 1,3%, niveau 5: 2,4%). Les degrés de difficulté objective des niveaux 1 à 5, tels qu'estimés à l'aide du modèle de Rasch, sont les suivants: -0,02, 1,18, 1,69, 1,78, 1,94.

Figure IV.18: Exemple d'exercice très difficile de compréhension générale de l'écrit en format ouvert (codage graduel à cinq niveaux)

<p>Question 8.1</p> <p>Quel titre pourrait-on donner à cette étude?</p> <p>«Selon la théorie d'Hippocrate, l'être humain se compose essentiellement de quatre humeurs; le sang, la lymphe, l'atrabile et la bile. Quand ces humeurs sont en équilibre, l'être humain est en bonne santé, et quand un déséquilibre intervient, l'être humain est malade. Quand quelqu'un avait de la fièvre, ce qui est très souvent le cas lors d'infections, comme nous le savons aujourd'hui, il avait, selon la théorie d'Hippocrate, trop de sang. Pour guérir le patient de sa fièvre, le traitement logique était de diminuer la quantité de sang, donc de lui administrer une saignée. Cette théorie prévalut pendant très longtemps, et c'est Robert Koch, un chercheur de Berlin, qui découvrit que les maladies infectieuses n'étaient pas dues à un excès de sang, mais à la présence de bactéries, notamment.</p> <p>Une infection, la tuberculose par exemple, n'est donc pas explicable par un déséquilibre des humeurs corporelles, mais par la présence dans le corps d'un agent pathogène qui provoque une atteinte des poumons, entraînant l'apparition de la fièvre. Cette théorie est appelée la théorie des agents pathogènes.</p> <p>Un autre exemple de théorie est celle des récepteurs. Le diabète, <i>Diabetes mellitus</i>, est une maladie connue depuis longtemps déjà; autrefois elle était nommée «urine sucrée». Au XIX^e siècle, Frerichs, un pathologiste berlinois, observa chez des patients atteints du Diabetes mellitus que leur pancréas, observé au microscope, avait une autre apparence que celui des personnes en bonne santé. Il en conclut que le diabète était une maladie du pancréas. Claude Bernard, le célèbre physiologiste français et fondateur de la médecine expérimentale, démontra que le sucre était produit dans le foie, et ainsi le diabète fut considéré comme une maladie du foie. En 1869, Langerhans décrivit pour la première fois les îlots de cellules situées dans le pancréas auxquels il donna plus tard son nom. Avec d'autres chercheurs il émit l'hypothèse que ces cellules régulaient le métabolisme du sucre. (...) En 1921, deux chercheurs canadiens, Bantig et Best, réussirent à isoler les îlots de cellules. (...) Ils injectèrent une substance extraite de cellules des îlots de Langerhans à un chien qui était en train de mourir du diabète. (...) Après avoir effectué des expériences sur eux-mêmes, ils injectèrent avec succès ces substances à des patients souffrant du diabète, et par la suite l'insuline devint une thérapie qui sauva la vie de beaucoup d'êtres humains, C'est ainsi que l'on apprit que l'insuline régula la teneur de sucre du sang.»</p> <p>(Steurer, 2005. <i>Wissenschaftstheoretische Grundlagen der Medizin</i>, note de cours en ligne OLAT, p. 7–9)</p> <p>Réponse:</p> <p>.....</p> <p>.....</p>	<p>F_A_81</p>
---	---------------

La question ci-après (figure IV.19) se réfère au même texte que le précédent. Il s'agissait de trouver un sous-titre adéquat pour chacun des trois paragraphes. La distribution des fréquences sur les différents niveaux de l'échelle est analogue à celle de l'exercice précédent, bien que l'on ait utilisé dans le cas présent une échelle à quatre niveaux seulement (contre cinq précédemment).

Figure IV.19: Exemple d'exercice difficile de compréhension générale de l'écrit en format ouvert
(codage graduel à quatre niveaux)

Question 8.2

Donnez un sous-titre approprié à chaque paragraphe.

a	F_A_82a
	
b	F_A_82b
	
c	F_A_82c
	

Les élèves ont été plus à l'aise avec cet exercice qu'avec le précédent. Ce n'est pas une surprise: dans le premier exercice, il fallait être capable de comprendre tous les paragraphes et de les mettre en relation. La difficulté objective δ des items calculée selon le modèle de Rasch s'établit comme suit pour les niveaux 1 à 4:

1^{er} paragraphe: 1,12, 1,07, 1,96, 2,18; 2^e paragraphe: -1,85, 1,77, 2,51, 2,59; 3^e paragraphe: -1,49, 1,25, 2,38, 2,45. Le premier et le dernier paragraphes sont les plus longs et les plus complexes. La qualité des sous-titres donnés par les élèves se distribue du reste de manière très similaire pour ces deux paragraphes. Les bonnes réponses de niveau 1 (réponse simple) représentent 52,4% du total au 1^{er} paragraphe et 58,3% au 3^e. Le taux de bonnes réponses de niveau 2 (réponses plus élaborées) ne dépasse pas 25,7% au 1^{er} paragraphe et 27,7% au 3^e. Le 2^e paragraphe n'est lui composé que de quatre phrases. 72,4% des élèves ont placé un titre simple, souvent constitué d'une notion tirée du texte (par exemple: «agents pathogènes», «théorie des agents pathogènes», «tuberculose»). Comme lors de l'exercice précédent, très peu d'élèves ont intégré dans leurs sous-titres des éléments dénotant des connaissances supplémentaires (de 0,4% à 3% selon le paragraphe) ou une certaine créativité sur le plan linguistique (de 0,9% à 3,2% selon le paragraphe).

Une proportion notable des élèves n'a pas été capable de fournir une proposition répondant aux exigences minimales. Pour le titre principal, 23,6% ont proposé un titre faux³² ou insuffisant. Pour les sous-titres, cette proportion varie entre 13,7% (paragraphe sur l'Antiquité) et 7,8% (paragraphe court sur la théorie des agents pathogènes).

La figure IV.20 donne un exemple d'item de difficulté moyenne qu'environ la moitié des bacheliers et bacheliers (54%) ont résolu correctement ($\delta = 0,27$). Il s'agissait de trouver la bonne définition du mot «fluorescent» à la lumière d'un petit texte traitant de mycologie où ce terme apparaît³³.

³² Les élèves n'ont pas reçu de points lorsque le titre principal proposé dénaturait le sens général du texte pour l'une des raisons suivantes: a) parce qu'il ne faisait référence qu'au contenu d'un seul paragraphe, b) parce qu'il ne restituait pas le propos essentiel du texte, c) parce qu'il n'évoquait qu'un des quatre aspects importants du texte (évolution, médecine, théorie et recherche, thérapie) ou d) mentionnait au moins deux éléments corrects mais y ajoutait un troisième qui était faux ou inapproprié.

Voici quelques exemples de propositions de surtitres qui n'ont pas reçu le moindre point: «Théories réfutées», «L'homme est guidé par les hormones», «Histoire des infections».

³³ Note du traducteur: le texte en question (texte 6.1.f) est reproduit dans la version allemande du présent rapport (figure IV.20).

Question 6.2

Que signifie «fluorescent» dans le texte 6.1.f?

F_W_62

Cochez la bonne réponse.

- 1 ☐ qui émet du fluor
 2 ☐ qui stocke du fluor
 3 ☐ qui éclaire de lui-même
 4 ☐ qui reflète la lumière

31,1% des élèves ont coché la réponse 4 «qui reflète la lumière». Cela laisse supposer qu'ils ne connaissaient pas l'adjectif «fluorescent» ni les verbes «absorber» et «émettre»³⁴. Ces derniers apparaissent en effet tous deux dans le texte de base de mycologie, où ils sont clairement associés à la fluorescence.

5.3 Résultats de mathématiques

5.3.1 Résultats d'ensemble

Le tableau IV.7 montre les résultats d'ensemble obtenus au test de mathématiques.

Tableau IV.7: Résultats du test de mathématiques

Types de résultats	Moyenne A	Ecart-type A	Maximum A	Moyenne P	Ecart-type P	Minimum P	Maximum P
TOT_MATH	37,65	13,25	80,00	-.09	.721	-2,22	2,44
ANALYSE	7,79	4,71	20,00	-.12	.896	-2,70	2,90
GEOMETRIE	9,08	4,58	20,00	-.08	.774	-2,28	2,49
STOCHASTIQUE	11,02	5,12	20,00	-.08	.990	-3,46	3,24
DIVERS_M	9,76	4,76	20,00	-.11	.818	-2,55	2,64

N = 13 524 (pondérés)

Explication des nouvelles abréviations apparaissant dans le tableau:

- TOT_MATH: résultat global en mathématiques
- ANALYSE: calcul analytique (fonctions, calcul infinitésimal, suites et séries)
- GEOMETRIE: géométrie (trigonométrie, géométrie analytique, calcul vectoriel)
- STOCHASTIQUE: stochastique (calcul des probabilités, analyse combinatoire, statistique)
- DIVERS_M: divers (théorie des ensembles/mathématiques discrètes, logique, algèbre élémentaire)
- Moyenne A: moyenne des points obtenus selon la méthode de calcul A (maximum = 20 points par domaine, soit 80 points en tout)

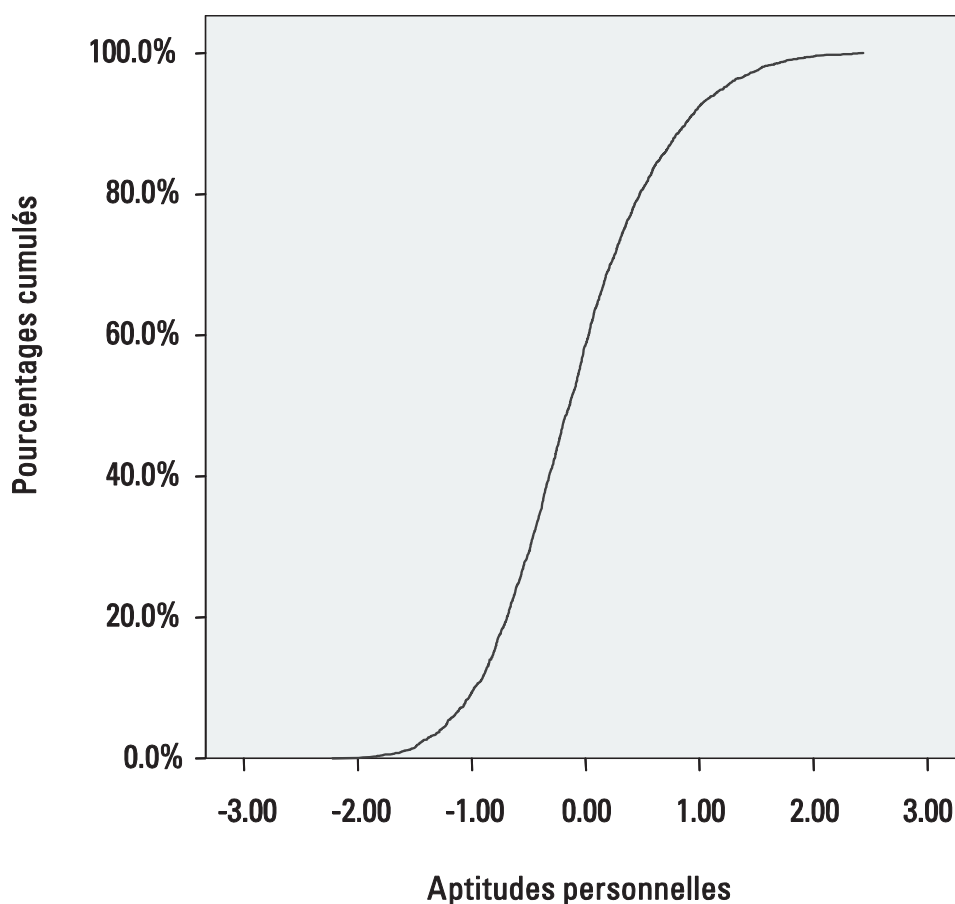
En moyenne, les élèves ont résolus correctement un peu moins de la moitié des items. Le taux de bonnes réponses se monte à environ 40% en calcul analytique, à 45% en géométrie, à 55% en stochastique et à 50% dans les autres domaines. Les écarts sont toutefois moindres lorsque l'on tient compte du degré de difficulté des items

³⁴ Note du traducteur: Une dimension particulière a pu jouer un rôle en Suisse alémanique. Comme le fait remarquer l'auteur du rapport, les mots allemands utilisés en lieu et place de «fluorescent» (*autofluoreszent*), «absorber» (*absorbieren*) et «émettre» (*emittieren*) n'ont pas de consonance germanique (ce sont des mots d'origine étrangère).

dans chacun de ces domaines. Les aptitudes estimées des élèves en stochastique par exemple – les plus élevées – se rapprocheraient alors de la moyenne. Le résultat brut relativement médiocre en calcul analytique doit être relativisé dans l'autre sens. Les items soumis aux élèves dans ce domaine étaient particulièrement difficiles. En moyenne, les aptitudes personnelles estimées des bacheliers et des bachelères sont telles qu'elles leur permettent de résoudre des items d'un niveau d'exigence légèrement inférieur à la moyenne dans tous les domaines de compétence. Les aptitudes personnelles en mathématiques sont plus petites que celles en langue première. Mais elles sont plus grandes que celles en biologie (en ne tenant compte dans ce dernier cas que des bacheliers et bachelères qui avaient terminé leurs cours dans cette discipline au moins six mois avant de passer le test). La dispersion des résultats des élèves est de nouveau considérable. Au moins un élève a obtenu le maximum des points possibles dans chaque domaine. La dispersion des résultats des classes est notable aussi. La meilleure classe (64,04 points) a résolu correctement près de trois fois plus d'items que la plus mauvaise (22,36 points).

La figure IV.21 représente graphiquement la distribution cumulée des aptitudes personnelles aux mathématiques estimées globalement.

Figure IV.21: Estimation globale des aptitudes personnelles en mathématiques



La fréquence d'apparition des différentes réponses dans les copies rendues par les élèves a été analysée en détail pour chaque exercice. Les résultats de cet examen doivent être interprétés avec prudence car le degré de difficulté des items varie. Les élèves ont résolu avec une certaine facilité les items portant sur des fonctions à deux variables, les additions à plusieurs indices et certaines questions de logique (tables de vérité). Les résultats sont très bons également en calcul vectoriel et en analyse combinatoire. Certains items de logique ont été particulièrement mal résolus (voir section 5.3.2). Dans une mesure moindre, les élèves ont éprouvé des difficultés avec les fonctions exponentielles et logarithmiques. Les résultats sont inférieurs à la moyenne

également pour les items portant sur les systèmes d'équations, les équations quadratiques et les problèmes de suites et de séries. Ceci pouvant expliquer cela, l'enseignement de ce dernier domaine tend à être délaissé dans les gymnases.

Les élèves ont très bien résolu les items mettant à contribution leur aptitude à comprendre un énoncé (qui est l'une des dimensions de la grille de compétences en mathématiques). Ils sont beaucoup plus à l'aise en calcul avec des chiffres (résultats supérieurs à la moyenne) que dans le maniement d'entités abstraites (variables ou inconnues). Les items d'interprétation de diagrammes ont été assez bien résolus aussi en moyenne.

5.3.2 Résultats d'items particuliers

Cette section s'intéresse à la manière dont une série d'items ont été plus ou moins bien résolus. Comme dans la section consacrée à la langue première, le taux de bonnes réponses sera mis en relation avec le degré de difficulté objectif de l'exercice (δ), estimé à l'aide du modèle de Rasch.

L'exercice le plus facile était un exercice de logique (voir figure IV.22). Il consistait à remplir une table de vérité composée de trois colonnes, constituant chacun un sous-exercice. Le maniement de ce type de tables n'est probablement plus guère enseigné au gymnase. Pour l'élève, il était donc important de bien comprendre l'énoncé. La première colonne a été remplie correctement par 91% des élèves ($\delta = -2,59$). C'est le taux le plus élevé à avoir été enregistré dans le cadre du test de mathématiques. Le taux de bonnes réponses s'est monté à 84% dans la 2^e colonne ($\delta = -0,98$) et à 83% dans la 3^e ($\delta = -1,89$).

Figure IV.22: Exercice de mathématiques facilement résolu (exemple 1)

Voici un extrait d'une banque de données qui recense des étudiants:

Numéro	Nom	Sexe	Pays d'origine	Faculté
1	Anna	féminin	France	Architecture
2	Sergio	masculin	Italie	Économie
3	Miriam	féminin	Suisse	Communication
4	Luc	masculin	France	Économie

Sachant que

A: "étudie l'économie"

B: "est de nationalité suisse"

C: "est une femme"

Remplissez la table de vérité suivante:

Numéro	Nom	A	A ou B	B et C
1	Anna	f		
2	Sergio	v		
3	Miriam			
4	Luc			

v signifie vrai, f signifie faux

M_X_10

L'exercice de calcul vectoriel reproduit dans la figure IV.23 a été très bien résolu également. Il implique de savoir interpréter des forces représentées sous forme de vecteurs et met aussi assez fortement à contribution la capacité de comprendre un énoncé.

La grande majorité des élèves a deviné que les deux décompositions vectorielles à gauche de la page sont correctes. Le taux de bonne réponses se monte à 72% pour la figure du haut ($\delta = -1,13$) et 89% pour celle du bas ($\delta = -2,36$). 79% des bacheliers et bacheliers ont senti que la figure en haut à droite est erronée ($\delta = -1,57$). La décomposition vectorielle figurant en bas à droite leur a causé un peu plus de problèmes. 61% des élèves ont vu qu'elle est correcte ($\delta = -0,59$).

Figure IV.23: Exercice de mathématiques facilement résolu (exemple 2)

En physique, il est souvent nécessaire de décomposer un vecteur \vec{a} en une somme de deux vecteurs \vec{a}_1 et \vec{a}_2 , dont les directions sont déterminées par des raisons physiques.

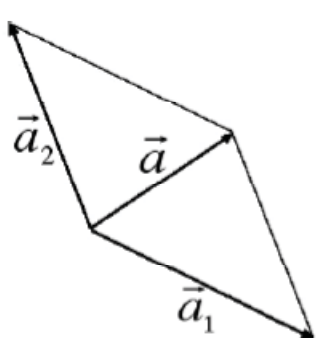
On a donc:

$$\vec{a} = \vec{a}_1 + \vec{a}_2$$

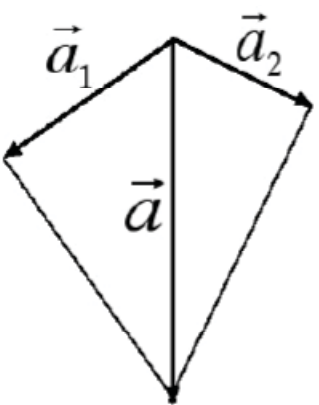
\vec{a}_1 et \vec{a}_2 sont les composantes du vecteur \vec{a} selon la direction 1 et la direction 2.

Lesquelles des décompositions sont correctes?

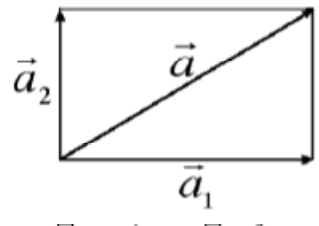
Cochez les bonnes réponses:



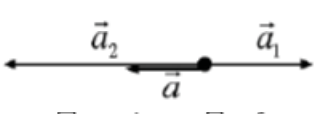
☐ vrai ☐ faux



☐ vrai ☐ faux



☐ vrai ☐ faux



☐ vrai ☐ faux

M_G_17

Les élèves n'ont pas éprouvé de grandes difficultés à résoudre l'exercice exposé dans la figure IV.24, ce qui était attendu. 88% ont trouvé la valeur correcte de p ($\delta = -2,23$) et 86% celle de q ($\delta = -2,07$).

164

Figure IV.24: Exercice de mathématiques facilement résolu (exemple 3)

Dans le diagramme ci-dessous, deux probabilités ne sont pas indiquées.
Complétez-les:

$p = \dots\dots\dots$

$q = \dots\dots\dots$

M_S_15

Les élèves s'en sont très bien sortis également sur l'exercice d'algèbre élémentaire qui suit (figure IV.25).

Figure IV.25: Exercice de mathématiques facilement résolu (exemple 4)

Quels couples $(m; n)$ satisfont l'inégalité

$$\frac{1}{m} + \frac{1}{n} > \frac{1}{2} ?$$

Cochez les bonnes réponses:

$(m; n) = (3; 4)$	<input type="checkbox"/> oui	<input type="checkbox"/> non
$(m; n) = (5; 3)$	<input type="checkbox"/> oui	<input type="checkbox"/> non
$(m; n) = (4; 4)$	<input type="checkbox"/> oui	<input type="checkbox"/> non
$(m; n) = (3; 3)$	<input type="checkbox"/> oui	<input type="checkbox"/> non

M_X_28

La fréquence des réponses données est la suivante:

Première proposition: 79% de réponses correctes ($\delta = -1,49$)

Deuxième proposition: 69% de réponses correctes ($\delta = -0,97$)

Troisième proposition: 85% de réponses correctes ($\delta = -1,98$)

Quatrième proposition: 76% de réponses correctes ($\delta = -1,36$).

L'exercice d'analyse combinatoire présenté à figure IV.26 a été résolu correctement par plus de 80% des élèves ($\delta = -1,63$). Il teste les connaissances des élèves en définitions et en propositions mathématiques.

Figure IV.26: Exercice de mathématiques facilement résolu (exemple 5)

<p>Un dé (usuel, à 6 faces) est lancé n fois de suite.</p> <p>Combien d'issues différentes cette expérience aléatoire a-t-elle?</p> <p>Cochez la bonne réponse:</p> <p><input type="checkbox"/> 1 n</p> <p><input type="checkbox"/> 2 2^n</p> <p><input type="checkbox"/> 3 $6n$</p> <p><input type="checkbox"/> 4 6^n</p>	M_S_02
---	--------

Un premier exemple d'exercice mal résolu figure dans la figure IV.27. Les élèves ont manifestement eu des problèmes à comprendre le concept de «dénombrabilité». Seuls 9% d'entre eux ont trouvé la bonne réponse ($\delta = 2,48$).

Figure IV.27: Exercice de mathématiques ayant posé problème (exemple 1)

<p>Un ensemble A est dénombrable, si:</p> <p><input type="checkbox"/> 1 A est fini</p> <p><input type="checkbox"/> 2 A possède une infinité d'éléments</p> <p><input type="checkbox"/> 3 il existe une bijection $\mathbb{N} \rightarrow A$</p> <p><input type="checkbox"/> 4 il existe une bijection $\mathbb{R} \rightarrow A$</p>	M_X_23
---	--------

L'exercice suivant (figure IV.28) n'a également été résolu que par 9% des bacheliers et bacheliers ($\delta = -2,48$). Ceux-ci ont probablement été perturbés par la désignation « $4\mathbb{Z}$ ».

Figure IV.28: Exercice de mathématiques ayant posé problème (exemple 2)

<p>Soit l'ensemble $4\mathbb{Z} = \{m \in \mathbb{Z} \mid m = 4z \text{ pour } z \in \mathbb{Z}\}$.</p> <p>Notez tous les nombres n qui appartiennent à l'ensemble $4\mathbb{Z}$ et pour lesquels on a $-10 < n < 10$.</p> <p>Réponse:</p>	M_X_19
--	--------

Les élèves n'ont été que 10% également à trouver la bonne réponse à l'exercice présenté dans la figure IV.29 ($\delta = 2,38$).

Figure IV.29: Exercice de mathématiques ayant posé problème (exemple 3)

<p>Pour des séries géométriques, on connaît la formule suivante:</p> $\sum_{k=0}^{\infty} a^k = a^0 + a^1 + a^2 + a^3 + \dots = \frac{1}{1-a} \text{ avec } a < 1$ <p>Appliquez cette formule à l'expression suivante:</p> $p \sum_{j=0}^{\infty} (1-p)^j \text{ avec } 1-p < 1$ <p>et simplifiez au maximum le résultat.</p> <p>Réponse: La somme vaut</p>	M_A_11
---	--------

La question reproduit dans la figure IV.30 comprend quatre sous-questions qui ont causé aux élèves des difficultés d'intensité très variable. Le taux de bonnes réponses est de 84% à la première question ($\delta = -18,87$), 49% à la deuxième question ($\delta = -0,02$), 20% à la troisième question ($\delta = 1,48$) et 11% seulement à la dernière question ($\delta = 2,23$).

Pour le reste, les résultats les moins homogènes entre élèves, et de loin, ont été enregistrés dans le cadre des items de logique et de ceux basés sur la théorie des ensembles. Certains items étaient très difficiles car ils contenaient des concepts que beaucoup d'élèves ne connaissaient manifestement pas.

Figure IV.30: Exercice de mathématiques avec des composants de difficulté variable (exemple)

<p>Parmi les fonctions suivantes:</p> $f(x) = x^3$ $g(x) = \frac{1}{x}$ $h(x) = -\frac{1}{x^2}$ $j(x) = \frac{2}{x}$ $k(x) = \ln(x^2)$ $l(x) = 3x^2$ $m(x) = \frac{1}{2\sqrt{x}}$ $n(x) = \ln(2x)$ $p(x) = 2^x$ <p>il y en a quatre qui sont des dérivées d'autres fonctions de cette liste.</p> <p>Complétez:</p> <p>La fonction $l(x)$ est la dérivée de la fonction</p> <p>La fonction est la dérivée de la fonction $g(x)$.</p> <p>La fonction $j(x)$ est la dérivée de la fonction</p> <p>La fonction est la dérivée de la fonction $n(x)$.</p>	M_A_06
--	--------

5.4 Résultats de biologie

5.4.1 Résultats d'ensemble

Le tableau IV.8 montre les résultats d'ensemble obtenus au test de biologie. Dans beaucoup de gymnases, les cours de biologie se terminent un à deux ans avant les examens de maturité. Les résultats sont présentés pour les deux catégories d'élèves («terminé oui», «terminé non»).

Tableau IV.8: Résultats du test de biologie

terminé	Moyenne A		Ecart-type A		Maximum A		Moyenne P		Ecart-type P		Minimum P		Maximum P	
	oui	non	oui	non	oui	non	oui	non	oui	non	oui	non	oui	non
TOT_BIO	26,8	32,5	8,71	8,84	54,2	56,3	−22	.10	.470	.474	−1,69	−1,28	1,24	1,59
B_EVOL	4,79	5,45	3,00	2,94	10,0	10,0	−21	.09	.452	.472	−1,68	−1,68	1,21	1,35
B_STRU	3,85	4,75	2,55	2,79	10,0	10,0	−22	.10	.460	.484	−2,07	−1,80	1,23	1,56
B_GEN	4,44	5,42	2,38	2,39	10,0	10,0	−21	.08	.410	.428	−1,81	−1,70	1,19	1,52
B_INFO	4,16	5,28	2,46	2,48	10,0	10,0	−24	.11	.497	.514	−1,97	−1,91	1,30	1,45
B_ECO	4,85	5,77	2,63	2,60	10,0	10,0	−24	.12	.533	.539	−2,22	−1,87	1,33	1,61
B_MET	4,56	5,78	2,53	2,56	10,0	10,0	−25	.14	.552	.569	−2,40	−2,07	1,52	1,76

N «terminé» = 9 693; N «pas terminé» = 3 748 (pondérés)

Explication des nouvelles abréviations apparaissant dans le tableau:

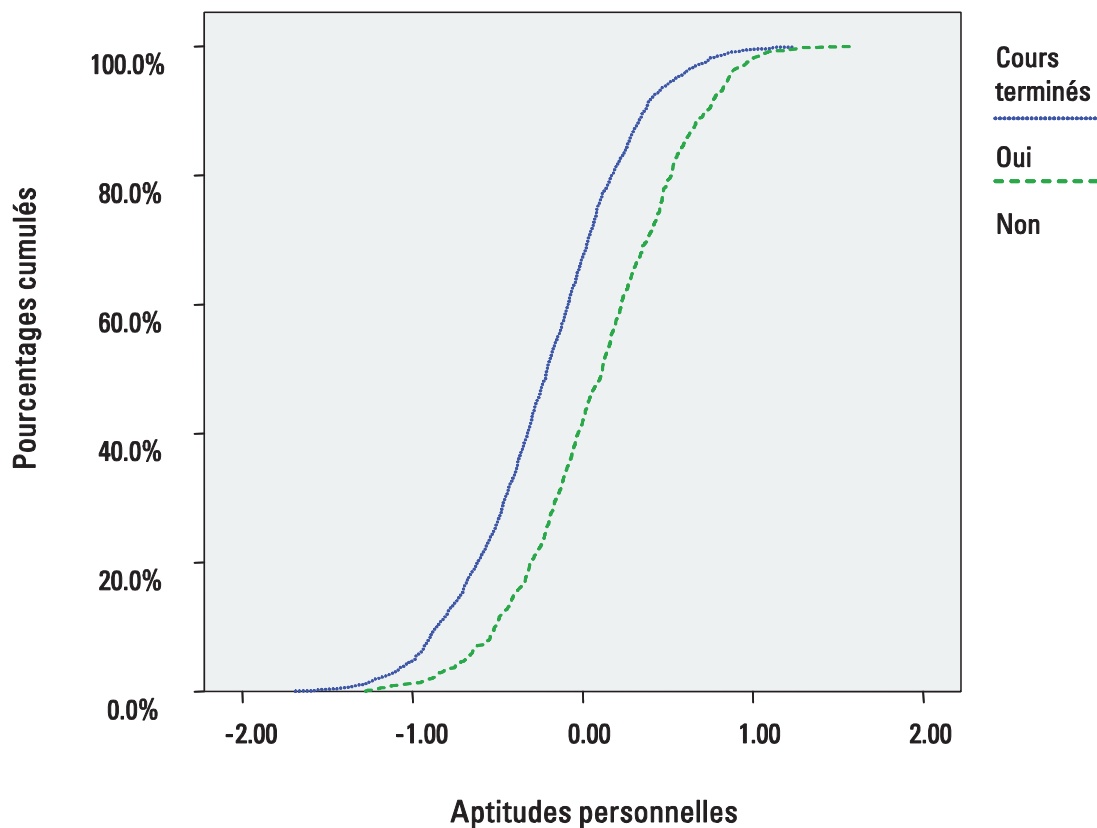
- TOT_BIOL: résultat global en biologie
- B_EVOL: évolution (processus de développement, évolution et projection)
- B_STRU: structure et fonction (construction et fonction des cellules, tissus et organes; différenciations fonctionnelles)
- B_GEN: génétique (bases de génétique moléculaire, application des connaissances et méthodes de la biologie moderne)
- B_INFO: traitement de l'information (communication entre cellules, traitement neuronal de l'information, perception)
- B_ECO: écologie (approche holistique, écologie et développement durable)
- B_MET: métabolisme (transformations moléculaires et énergétiques: nécessité et voies de transformation énergétique, anabolisme et catabolisme relatifs aux structures cellulaires et de l'organisme; régulation moléculaire du métabolisme: enzymologie)
- Moyenne A: moyenne des points obtenus selon la méthode de calcul A (maximum = 10 points par domaine, soit 60 points en tout)

En moyenne, près de la moitié des items de biologie ont été résolus correctement. Le groupe d'élèves qui avaient terminé les cours de biologie depuis un certain temps a obtenu des résultats tendanciellement moins bons. Leurs résultats en biologie sont plutôt plus mauvais que ceux qu'ils ont obtenus en langue première et en mathématiques. Leurs aptitudes personnelles estimées sont inférieures à un niveau d'aptitudes moyen. L'oubli progressif des connaissances acquises joue sans doute un rôle, mais on pouvait s'attendre à ce que cet effet soit encore beaucoup plus prononcé. La dispersion des résultats est à nouveau considérable tant au niveau des élèves que des classes. La meilleure classe a de nouveau résolu correctement presque trois fois plus d'items que la plus mauvaise (la moyenne varie entre 14,93 et 42,35). Personne n'a obtenu le nombre maximum de points possibles

sur l'ensemble du test, mais certains s'en sont beaucoup approchés. Au moins un élève a rendu une copie parfaite dans chaque domaine du test.

La figure IV.31 représente graphiquement la distribution cumulée des aptitudes personnelles à la biologie, estimées globalement.

Figure IV.31: Estimation globale des aptitudes personnelles à la biologie



5.4.2 Résultats d'items particuliers

Cette section s'intéresse à la manière dont une série d'items ont été plus ou moins bien résolus. Comme dans les parties consacrées à la langue première et aux mathématiques, le taux de bonnes réponses sera mis en relation avec le degré de difficulté objectif de l'exercice (δ), estimé à l'aide du modèle de Rasch.

Figure IV.32: Item facile de biologie (exemple 1)

(FKT_03)

De quels composants les protéines sont-elles constituées?

- ☐ De nucléotides
- ☐ De gènes
- ☐ De molécules de glucose
- ☐ D'acides aminés

La probabilité de résoudre l'item reproduit dans la figure IV.32 est de 0,857 ($\delta = -2,02$). Sur 100 étudiantes et étudiants venant d'entrer à l'université, 86 seraient donc vraisemblablement capables de faire le lien entre la chaîne fonctionnelle qu'est une protéine, et ses éléments constitutifs que sont les acides aminés. Mais ils seraient tout de même 14 à ne pas y parvenir.

Figure IV.33: Item facile de biologie (exemple 2)

(INF_07)

**Le taxol, issu de l'if, est un inhibiteur de la mitose et est employé dans la thérapie contre le cancer.
Comment agit le taxol?**

- ☐ Le taxol inhibe la mort cellulaire des cellules du corps.
- ☐ Le taxol inhibe la formation de cellules sexuelles.
- ☐ Le taxol inhibe l'action du système immunitaire.
- ☐ Le taxol inhibe la prolifération des cellules cancéreuses.

La phrase qui introduit la question dans l'exercice reproduit dans la figure IV.33 est extraite d'une note de cours de l'Université de Zurich («*Biodiversität: Pflanzen, Wirbeltiere, Pilze* » de Peter Linder, p. 72). Il est possible que la référence à la thérapie contre le cancer en relation avec l'unité de sens «inhibiteur de la mitose» aide à choisir la bonne solution, le cancer étant étroitement associé à l'idée de prolifération des cellules même dans le langage courant. Cet effet facilitateur agirait toutefois exactement de la même manière en situation réelle de lecture pendant les études. C'est pourquoi on décida d'en tenir compte et même d'en tirer volontairement parti lors de la conception de l'item, en estimant que le bachelier ou la bachelière qui interpréterait de façon erronée l'unité de sens «inhibiteur de la mitose» malgré cet indice (en l'associant en l'occurrence à la mort cellulaire, aux cellules sexuelles ou au système immunitaire) ne comprendrait pas non plus correctement la notion en lisant la note de cours à l'université. La probabilité de résoudre l'item se monte à 0,809 ($\delta = -1,65$). En moyenne, 19 nouveaux étudiants sur 100 ne parviendraient donc pas à établir d'association entre la mitose (l'objet de la question) et la prolifération de cellules cancéreuses (la bonne réponse) malgré le lien établi la mitose et la thérapie contre le cancer. Cet item s'est avéré sélectif (discrimination $> 0,25$) bien qu'il ne présente pas un grand degré de difficulté. Il a surtout été fatal à des bachelières et bacheliers qui s'avèrent faibles en biologie en général d'après leurs résultats au test.

Figure IV.34: Item facile de biologie (exemple 3)

(OEK_03)

**Un parasite nuit à un autre organisme, mais tire un avantage de cette cohabitation.
Quelle précision peut-on encore ajouter?**

- ☐ Il vit sur ou dans l'autre organisme.
- ☐ Il vit sur l'autre organisme.
- ☐ Il vit dans l'autre organisme.
- ☐ Il ne touche pas directement l'autre organisme.

77 bachelières et bacheliers sur 100 semblent savoir que les parasites peuvent vivre non seulement sur, mais aussi à l'intérieur de leur hôte. C'est en tous cas la proportion des élèves qui ont résolu correctement l'exercice

reproduit dans la figure IV.34 ($\delta = -1,43$). La famille des endoparasites comprend notamment le virus de la maladie du sommeil et les organismes monocellulaires à l'origine de la malaria (plasmodium). Mais son représentant de loin le plus connu est le fameux ver solitaire. On peut imaginer que les élèves y ont pensé et que cela les a aidés. Mais il est possible aussi que le taux élevé de bonnes réponses à cet exercice traduise des connaissances en biologie plus étendues.

Figure IV.35: Item facile de biologie (exemple 4)

(STW_09)

Dans quel processus les chloroplastes jouent-ils un rôle central?

- ☐ Dans la division cellulaire
- ☐ Dans la respiration cellulaire
- ☐ Dans la production d'ATP
- ☐ Dans la photosynthèse

L'exercice représenté dans la figure IV.35 teste chez les bacheliers et bachelères des connaissances qui seront considérées comme acquises lorsqu'ils seront à l'université. 75,5% des élèves ont trouvé la bonne solution (les chloroplastes jouent un rôle dans la photosynthèse). La proposition 2 (les chloroplastes jouent un rôle dans la respiration cellulaire) a induit en erreur une certaine frange d'élèves, victimes d'une confusion très fréquente entre les processus d'assimilation et de dissimilation ($\delta = -1,32$).

Figure IV.36: Item de difficulté moyenne de biologie (exemple 1)

(STW_07)

«Tous les animaux doivent remplir les mêmes tâches vitales fondamentales, entre autres le contrôle de l'hydratation et de la concentration saline (osmorégulation)».

Quelles sont les deux propositions concernant le processus de l'osmose qui sont correctes?

- I. Les particules salines migrent à travers la membrane semi-perméable.
- II. Les molécules d'eau migrent à travers la membrane semi-perméable.
- III. La membrane semi-perméable correspond, chez les êtres vivants, à une membrane cellulaire.
- IV. Le passage à travers la membrane semi-perméable consomme de l'énergie (ATP).

- ☐ I et IV
- ☐ I et III
- ☐ II et III
- ☐ II et IV

Seuls 51,7% des bacheliers et bachelères ont coché la bonne réponse (proposition 3) dans l'exercice représenté dans la figure IV.36 ($\delta = -0,16$). Le taux de bonnes réponses est étonnamment bas. L'osmose joue en effet un rôle fondamental dans tous les processus physiologiques. Deux raisons peuvent l'expliquer:

- a) Faire comprendre le principe de l'osmose est l'une des tâches les plus délicates et laborieuses dans l'enseignement de la biologie. Le concept d'osmose est une représentation modélisée du comportement

de micro-organismes invisibles. Une grande capacité d'abstraction et d'imagination est nécessaire pour pouvoir l'appréhender.

- b) L'osmose et la diffusion font en général partie des premières notions enseignées en biologie au gymnase en raison de l'importance de ces phénomènes. Beaucoup d'élèves, mais aussi d'enseignantes et d'enseignants, tournent ensuite rapidement la page. Pourtant, la capacité de se représenter des choses abstraites et en particulier des micro-organismes inobservables ne se développe vraiment qu'au cours du gymnase. Un rafraîchissement des connaissances sur l'osmose et la diffusion n'intervient que rarement avant la maturité, malgré le risque que cela comporte que les élèves comprennent mal d'autres phénomènes qui y sont liés.

Figure IV.37: Item de difficulté moyenne de biologie (exemple 2)

(GEN_07)

Qu'est-ce qu'un zygote?

- ☐ Un zygote est un embryon avant l'implantation dans l'utérus.
- ☐ Un zygote est une cellule formée par la fusion de deux gamètes.
- ☐ Un zygote est un embryon au stade biocellulaire
- ☐ Un zygote est un ovule qui n'a pas été fécondé.

L'item représenté dans la figure IV.37 n'a été résolu correctement que par la moitié des élèves également (51,7%, avec $\delta = -0,20$).

On peut se demander ici dans quelle mesure certains enseignants, soucieux de ne pas surcharger inutilement les capacités cognitives de leurs élèves, renoncent systématiquement à employer des termes techniques comme «zygote» au détriment d'expressions plus explicites pour désigner la même chose («cellule fécondée»). Ce réflexe ne nuit pas en tant que tel à la capacité de comprendre d'autres phénomènes fondamentaux qui y sont liés (la fusion des gamètes mâle et femelle par exemple). Mais il est problématique si la suite des cours est organisée en partant de l'idée que le concept «zygote» est acquis. Repousser l'assimilation de concepts clés peut progressivement nuire à l'aptitude aux études et conduire à des difficultés ultérieures parfois importantes à l'université, si l'étudiant doit constamment aller vérifier le sens de termes spécialisés qu'il rencontre dans les supports de cours.

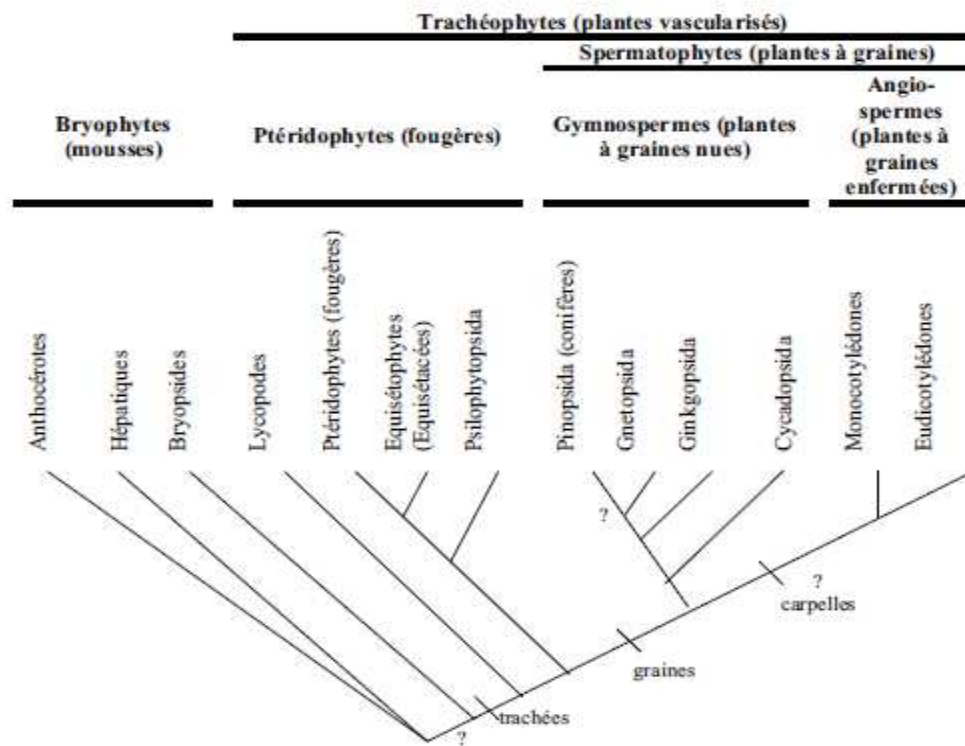
L'item reproduit dans la figure IV.38 met à l'épreuve la capacité des élèves à décrypter une arborescence ($\delta = -0,04$). A l'université, comprendre la langue symbolique de la phylogénie sera considéré comme acquis. Pourtant, dans un auditoire comprenant 100 nouveaux étudiants, 52 ne seraient pas en mesure de tirer des informations correctes de l'examen d'une arborescence phylogénétique. Seuls 48,1% des élèves ont trouvé la bonne réponse (proposition 2).

Figure IV.38: Item de difficulté moyenne de biologie (exemple 3)

(EVO_07)

Quelle proposition peut-on déduire du graphique suivant?

Phylogénie des embryophytes (plantes terrestres)



- ☐ Les conifères produisent des graines, mais pas les monocotylédones.
- ☐ Les conifères et les monocotylédones produisent des graines.
- ☐ L'ancêtre commun des fougères et des conifères produisait déjà des graines.
- ☐ Toutes les plantes terrestres actuelles produisent des trachées, des graines et des carpelles.

Figure IV.39: Item de difficulté moyenne de biologie (exemple 4)

(OEK_20)

Le *Prochlorococcus marinus* produit des cellules sphériques individuelles (0,6–0,8 µm de diamètre). Ce type d'algue est largement répandu dans le plancton des mers. Qu'est-ce que cela signifie?

- ☐ Ce type d'algue produit des cellules qui dérivent et flottent librement dans l'eau.
- ☐ Ce type d'algue ne vit que dans l'eau douce.
- ☐ Ce type d'algue vit à proximité du littoral sur le fond de la mer.
- ☐ Ce type d'algue produit de nouveaux nutriments à partir d'organismes morts.

L'item ci-dessus (figure IV.39) repose sur une note de cours de premier semestre universitaire dans laquelle apparaît le terme de «plancton». Plus de la moitié des élèves (52,8%) ont coché la bonne réponse ($\delta = -0,19$). Les autres ont établi un lien erroné avec le milieu naturel (proposition 2: eau douce; proposition 3: fond marin) ou avec le rôle dans l'écosystème (proposition 4: décomposeur au lieu de producteur).

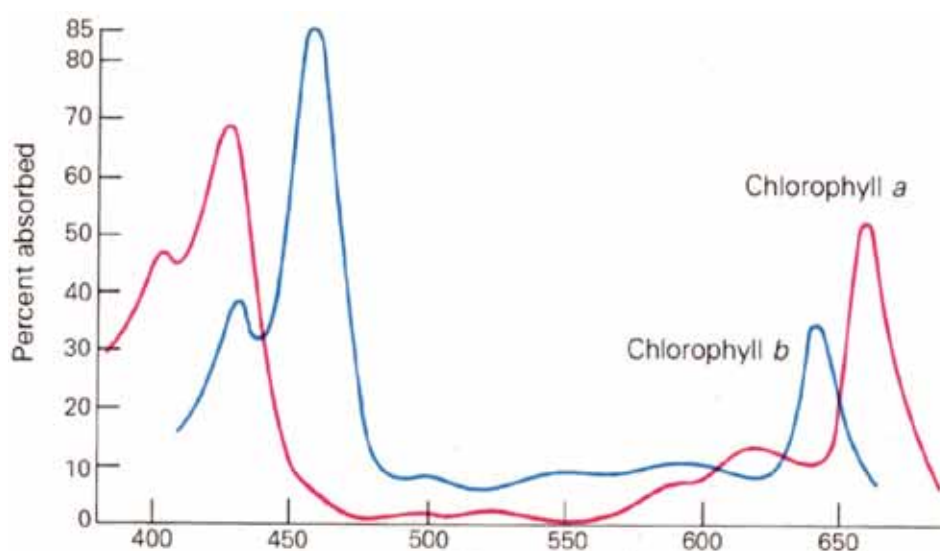
Les items suivants ont tous été plutôt mal résolus.

Figure IV.40: Item difficile de biologie (exemple 1)

(STW_21)

Dans le diagramme ci-dessous présente deux courbes dans un système de coordination. L'axe des x représente une échelle et une unité de grandeur typique, de même que l'axe des y représente une échelle et une unité de grandeur typique.

Comment s'appelle la propriété de la chlorophylle a ou de la chlorophylle b qui est toujours présentée de cette manière?



L'exercice reproduit dans la figure IV.40 est le premier dans cette partie où les élèves devaient se souvenir d'un terme technique sans que des solutions leur soient proposées.

Les expressions considérées comme une bonne réponse en français étaient «spectre d'absorption», «spectre d'absorption de l'énergie lumineuse» ou «capacité d'absorption du spectre lumineux», et en allemand «Absorptionsspektrum». Beaucoup d'élèves ont paru désemparés face à cette question. Seuls 10,7% des élèves ont trouvé la bonne réponse ($\delta = 1,78$). Beaucoup n'ont pas répondu du tout (44% en Suisse alémanique et 64% en Suisse romande). De nombreuses réponses fausses étaient très éloignées de la bonne solution. En Suisse romande, les réponses suivantes ont par exemple été données: couleur, photosynthèse, luminosité, pigmentation (en Suisse alémanique: *Aufnahmefähigkeit, Durchlässigkeit des Lichts, Energiegewinnung, Farbe, Lichtintensität, Photosynthese/Fotosynthese, Wirkungsspektrum*).

Figure IV.41: Item difficile de biologie (exemple 2)

(INF_01)

La rétine a une carte neuronale sur le corps géniculé latéral. Que signifie le terme «carte neuronale»?

- ☐ L'activité des cellules sensorielles dans la rétine est représentée sur une zone bidimensionnelle dans le corps géniculé latéral.
- ☐ L'activité des cellules sensorielles dans la rétine est enregistrée temporairement dans le corps géniculé latéral.
- ☐ L'activité de chaque cellule sensorielle de la rétine est dirigée de façon ciblée par le corps géniculé latéral.
- ☐ L'activité de chaque cellule du corps géniculé latéral est dirigée de façon ciblée par les cellules sensorielles de la rétine.

Seuls 22,5% des élèves alémaniques et 12,8% des élèves romands ont fourni la bonne réponse à l'exercice reproduit dans la figure IV.41 ($\delta = 1,41$). D'un côté comme de l'autre de la Sarine, ils s'en seraient probablement mieux sortis en répondant au hasard (une chance sur quatre de tomber juste). Une part prépondérante des bacheliers et bacheliers ne maîtrisent donc pas le sens de l'expression «carte neuronale» au moment d'entamer leurs études supérieures. Dans l'ensemble, les bacheliers et bacheliers ont jugé les mauvaises réponses proposées plus plausibles que la bonne (proposition 1). Ils ont été leurrés dans les mêmes proportions par les différentes mauvaises réponses, qu'ils soient faibles ou forts sur l'ensemble du test de biologie. Lorsque les élèves ne connaissaient pas «spontanément» la bonne réponse à cet item, ils ont apparemment tous essayé de deviner la bonne réponse de la même manière. L'exercice a donc une valeur discriminatoire quasi-nulle, et ce en Suisse alémanique comme en Suisse romande. Une part sur-proportionnelle des élèves n'a pas répondu du tout à cette question (16,3% en Suisse alémanique, 15,8% en Suisse romande). Cela confirme clairement qu'il s'agissait d'un exercice vraiment difficile.

Figure IV.42: Item difficile de biologie (exemple 3)

(FKT_13)

Les corps fructifères de certains protozoaires ont des ressemblances morphologiques avec ceux des champignons. De quel type de ressemblance s'agit-il?

- ☐ de la ressemblance du taux de reproduction
- ☐ de la ressemblance de l'histoire de l'évolution
- ☐ de la ressemblance du métabolisme dans les cellules du corps fructifère
- ☐ de la ressemblance de l'aspect

La situation est très similaire pour l'item exposé dans la figure IV.42 ($\delta = 1,23$). 24,2% des élèves ont trouvé la bonne réponse en Suisse alémanique et 17,7% en Suisse romande. Le taux de bonnes réponses est de nouveau inférieur à la probabilité de trouver la bonne réponse en se laissant guider par le hasard (25%). Les mauvaises réponses proposées aux élèves ont toutes une valeur discriminatoire très proche de zéro. Ceux qui ne connaissaient pas d'emblée la bonne réponse manquaient de repères pour essayer de la deviner malgré tout. Tant les bons que les mauvais élèves en biologie se sont laissés induire en erreur sur-proportionnellement par la proposition 3 (34,9% en Suisse alémanique, 43,0% en Suisse romande). Contrairement aux autres, cette proposition répète un concept qui apparaît dans l'énoncé de l'exercice («corps fructifère»). Un professeur d'université ne peut donc pas partir du principe qu'il peut utiliser sans autre le terme «morphologie» pour désigner la «forme et structure extérieure visible» d'un être vivant dans ses premiers cours. Dans la majorité des cas, ses nouveaux étudiants feront des associations erronées à partir de ces termes.

(FKT_01)

Quelles caractéristiques appartiennent à la description anatomique de l'appareil digestif d'une vache?

- I. le poids total de l'intestin grêle
- II. la longueur totale de l'œsophage
- III. la valeur du pH dans l'estomac
- IV. la part de poids des bactéries vivantes dans le tractus intestinal
- V. l'émission de méthane par heure
- VI. la partie la plus épaisse de la muqueuse de l'intestin grêle

- ☐ I, II, VI
- ☐ I, III, VI
- ☐ I, II, IV, V
- ☐ I, III, IV, V

L'item reproduit dans la figure IV.43 teste la capacité des élèves à distinguer les caractéristiques anatomiques et les caractéristiques physiologiques d'une vache ($\delta = 1,20$). Les premières se réfèrent à la forme et à la structure intérieure et extérieure du ruminant, tandis que les secondes renvoient aux processus qui découlent de son anatomie. L'affirmation III n'est pas une caractéristique anatomique car la valeur du pH n'est ni une forme, ni une structure. Il en va de même pour la part de poids des bactéries dans le tractus (affirmation IV) et pour le volume de méthane émis (affirmation V). Seules les affirmations I, II et VI font référence à des caractéristiques anatomiques. C'est donc la proposition 1 qu'il fallait cocher. Il est frappant que les bacheliers et bacheliers ont opté presque aussi souvent pour chacune des quatre réponses possibles. Le phénomène est particulièrement marqué en Suisse romande (proposition 1: 22%, proposition 2: 20%, proposition 3: 24%, proposition 4: 22%). La distribution des réponses est très homogène en Suisse alémanique aussi (proposition 1: 22%, proposition 2: 21%, proposition 3: 18%, proposition 4: 30%). 9% des élèves n'ont pas répondu à cette question en Suisse romande et 12% en Suisse alémanique. Une nouvelle fois, le taux de bonnes réponses n'a pas dépassé 25% pour cet exercice très difficile. Le sens du terme «anatomie» ne peut donc être considéré comme acquis au moment où les élèves entament leurs études supérieures.

5.5 Comparaison avec les notes de maturité

La grande dispersion des résultats des tests observée dans tous les domaines est surprenante. Les tests se sont en effet déroulés peu avant les examens de maturité. Or ce diplôme est censé attester que son titulaire a priori la maturité nécessaire pour entamer des études supérieures dans n'importe quelle branche. Cela ne semble pas être le cas. Les bacheliers et bachelères ne disposent manifestement pas tous, dans l'ensemble des domaines testés, des compétences requises pour aborder des études universitaires. Le tableau IV.9 ci-après permet de comparer la moyenne et à la dispersion des notes de maturité délivrées dans les trois disciplines testées dans le cadre d'EVAMAR II.

Tableau IV.9: Sélection de notes de maturité

Types d'examens	N (pondéré)	Minimum	Maximum	Moyenne	Ecart-type
Langue première: écrit	13'394	2,00	6,00	4,41	.739
Langue première: note finale	13'390	2,50	6,00	4,63	.542
Mathématiques: écrit	13'390	1,00	6,00	4,03	1,048
Mathématiques: note finale	13'381	1,50	6,00	4,34	.826
Biologie: notation continue	11'112	2,50	6,00	4,68	.541
Sciences expérimentales: note finale	13'369	2,50	6,00	4,60	.566

Globalement, la dispersion des notes de maturité s'avère nettement plus faible que celle des résultats des tests. Elle est nettement plus importante en mathématiques que dans les deux autres disciplines. Le constat vaut tout particulièrement pour l'examen écrit de mathématiques. C'est le seul examen où la note la plus basse a été attribuée (1) et où la note moyenne est juste suffisante (4,0 exactement). Dans les classes les plus faibles, les élèves ont réalisé les moyennes suivantes lors des différentes épreuves de maturité: écrit de langue première: 3,6; note finale de langue première: 3,97; écrit de mathématiques: 2,88; note finale de mathématiques: 3,73 ; notation continue en biologie (moyenne durant la dernière année de cours): 3,85; note finale de sciences expérimentales: 4,0. Dans les meilleures classes, les moyennes suivantes ont été enregistrées: écrit de langue première: 5,11; note finale de langue première: 5,17; écrit de mathématiques: 5,48 ; note finale de mathématiques: 5,44; notation continue en biologie: 5,28; note finale de sciences expérimentales: 5,31.

Les proportions d'élèves dont les performances ont été jugées insuffisantes par leur école ou leur enseignant dans les différents domaines sont les suivantes (note délivrée inférieure ou égale à 3,9):

- Ecrit de langue première: 19,6%
- Note finale de langue première: 4,7%
- Ecrit de mathématiques: 41,4%
- Note finale de mathématiques: 24,4%
- Notation continue en biologie: 5,6%
- Note finale de sciences expérimentales: 5,6%.

A l'écrit, les performances ont donc été jugées insuffisantes dans 41,4% des cas en mathématiques et dans 19,6% des cas en langue première (l'exercice consistait en général à rédiger une dissertation).

Les chiffres font un bond dès que l'on s'approche de la zone où les performances commencent à être jugées suffisantes comme le montre les données suivantes (1^{er} chiffre = proportion d'élèves ayant obtenu une note insuffisante; 2^e chiffre = proportion des élèves dont la note est comprise entre 3,9 et 4,0):

- Ecrit de langue première: bond de 19,6% à 37,9%
- Ecrit de mathématiques: bond de 41,4% à 56,7%
- Notation continue en biologie: bon de 5,6% à 16,7%.

On peut donc supposer que les examinatrices et examinateurs ont tendance à «arrondir» la note vers le haut lorsque les résultats d'un examen sont un peu insuffisants. La proportion des examens objectivement insuffisants est donc probablement encore plus élevée que cela ne ressort de ces chiffres.

Sur la base des résultats des tests, nous avons déjà émis ci-dessus l'hypothèse que les compétences des bacheliers et bacheliers sont en partie insuffisantes. Cela se reflète dans la pratique d'évaluation aux examens de maturité. Seuls les chiffres précis divergent. On peut penser par ailleurs que la manière d'attribuer des notes s'adapte en partie au niveau de la classe.

5.6 Résultats du test des compétences transversales (TCT)

Le tableau IV.10 montre les résultats obtenus dans les différents volets du test des compétences transversales. Comme cela a déjà été mentionné, le TCT n'a eu lieu qu'en Suisse alémanique. Les valeurs obtenues pour chacun des quatre domaines de compétence testés résultent de l'addition des résultats de chaque exercice. A noter que certains élèves n'ont pas participé aux items de tous les domaines de compétence, mais les cahiers de test ont été distribués de façon homogène.

Tableau IV.10: Résultat du test des compétences transversales (TCT)

Types de résultats	Moyenne A	Ecart-type A	Maximum A	Moyenne P	Ecart-type P	Minimum P	Maximum P
T_QUANT	4,57	2,496	10,00	.06	.839	-2,93	2,55
T_DIAGR	5,28	2,270	10,00	.08	.779	-2,37	2,33
T_TEXT	5,75	2,512	12,00	.09	.717	-2,39	2,32
T_NAT	4,58	2,191	10,00	.04	.712	-2,39	2,41

N = 9 281 (pondéré)

Explication des nouvelles abréviations apparaissant dans le tableau:

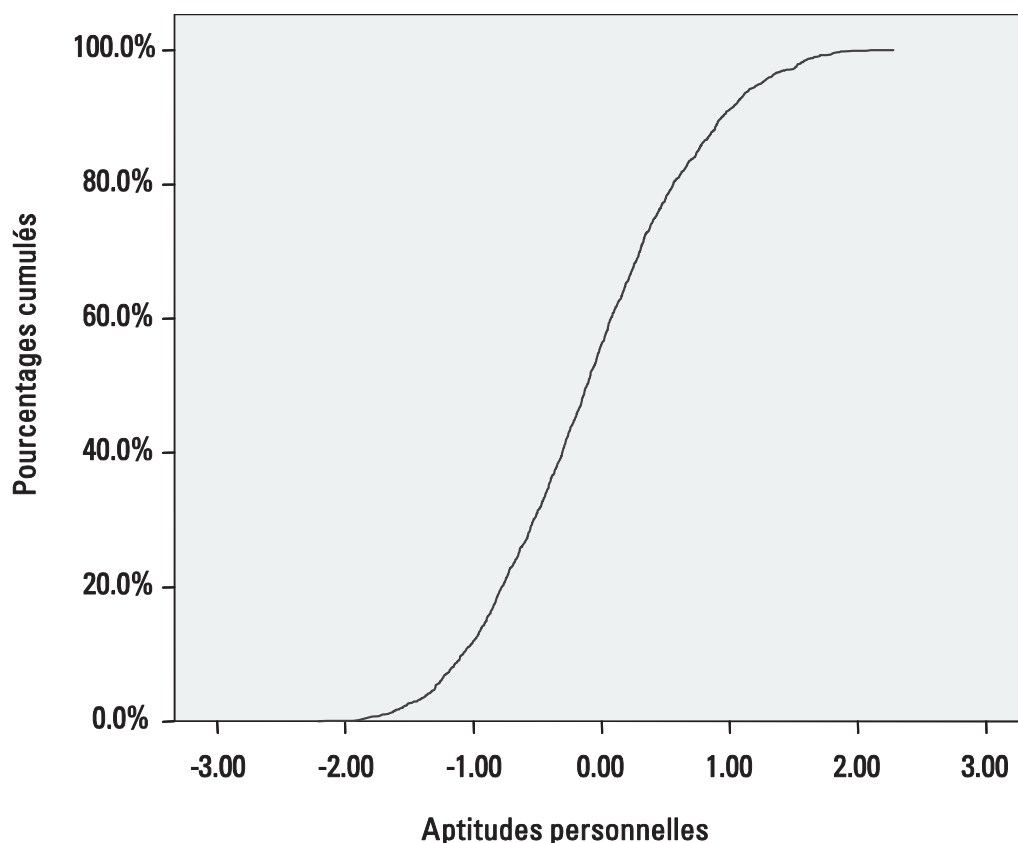
- T_QUANT: ensembles et formes
- T_DIAGR: diagrammes et tableaux
- T_TEXT: compréhension de textes
- T_SCIEN: notions élémentaires de sciences expérimentales
- Moyenne A: valeur moyenne selon la méthode de calcul A (maximum = 10 points par domaine, 12 points pour T_TEXT)

Les élèves ont résolu correctement un peu moins de la moitié des items du TCT dans les domaines «Ensembles et formes», «Compréhension de textes» et «Notions élémentaires de sciences expérimentales». Le taux de bonnes réponses ne dépasse la moitié que dans le domaine «Diagrammes et tableaux». En moyenne, les résultats sont légèrement moins bons que ceux enregistrés lors du test d'aptitudes pour les études de médecine (AMS). C'est logique: seuls les élèves qui le souhaitent participent à l'AMS. Une présélection a donc lieu sur une base volontaire. En fait, il est étonnant que l'écart entre les résultats ne soit pas plus important. Les élèves semblent donc s'être appliqués en remplissant le TCT. Ceci constitue un indice important que l'ensemble des tests d'EVAMAR II ont dans une large mesure été remplis sérieusement par les élèves. En moyenne, les aptitudes personnelles estimées des bacheliers et bacheliers sont telles qu'elles leur permettent de résoudre des items d'un niveau d'exigence légèrement supérieur à la moyenne dans tous les domaines de compétence. La dispersion des résultats est de nouveau frappante. Un élève au moins a rendu une copie parfaite dans chacun des domaines. La dispersion des résultats du TCT au niveau des classes est la plus importante de tous les tests. Les moyennes de la classe la plus faible sont les suivantes dans les différents domaines du TCT: 1,60 pour T_QUANT, 2,50 pour

T_DIAGR, 2,44 pour T_TEXT et 2,25 pour T_SCIEN. Les valeurs maximales sont les suivantes: 8,30 pour T_QUANT, 8,40 pour T_DIAGR, 9,56 pour T_TEXT et 7,67 pour T_SCIEN.

La figure IV.44 représente graphiquement la distribution cumulée des aptitudes personnelles estimées pour le TCT dans son ensemble.

Figure IV.44: Estimation globale des aptitudes personnelles d'après les résultats au TCT



5.7 Résultats du questionnaire complémentaire

Cette section expose quelques résultats du questionnaire complémentaire, en particulier les appréciations que les élèves ont portées sur les tests et sur leurs propres compétences transversales.

5.7.1 Perception des tests

En moyenne, les bacheliers et bacheliers ont trouvé les tests plutôt peu intéressants (1 = pas intéressant, 4 = très intéressant). Les résultats sont très voisins pour les tests de toutes les disciplines. La moyenne se situe à 2,15 pour les mathématiques (écart-type = 0,92), à 2,13 pour la biologie (écart-type = 0,85) et à 2,16 pour la langue première (écart-type 0,87). Les différences sont significatives entre les régions linguistiques dans les trois disciplines. Les élèves alémaniques ont été plus intéressés que les romands par le test de langue première (moyenne de 2,22 [écart-type = 0,88] contre 2,03 en Suisse romande [écart-type = 0,83]). Il en va de même pour le test de biologie (moyenne de 2,21 [écart-type = 0,96] contre 1,97 en Suisse romande [écart-type = 0,90]). En revanche, les élèves romands ont manifesté un intérêt plus marqué que les alémaniques pour les items de mathématiques (moyenne de 2,26 [écart-type = 0,91] contre 2,10 en Suisse alémanique [écart-type = 0,93]). Le test des compétences transversales – qui ne s'est déroulé qu'en Suisse alémanique – a été significativement mieux noté par les élèves. La moyenne s'élève à 2,39 (écart-type = 0,93), une valeur proche de la valeur moyenne théorique (2,5).

Les élèves semblent avoir participé aux tests avec application, malgré la distance affichée à leur égard. Une échelle a été construite à partir de trois items mesurant leur assiduité, leur concentration et leur engagement lors des tests. La moyenne s'élève à 2,87 (écart-type = 0,67), soit plus que la moyenne théorique (2,5). Dans l'ensemble, les élèves étaient donc motivés. Les valeurs sont comparables de part et d'autre de la Sarine.

5.7.2 Perception des compétences transversales propres

Diverses compétences transversales jugées importantes à l'école, notamment par le corps professoral (voir sous-projet A4), ont été testées dans le questionnaire complémentaire. Il s'agit de la gestion de la planification du temps de travail, de l'efficacité de la préparation des examens, de la gestion du temps, de la concentration, de la capacité d'analyse et de synthèse, de la capacité d'argumenter et de la capacité d'adopter d'autres points de vue que le sien (voir point 2.9). Les élèves étaient invités à dire dans quelle mesure diverses affirmations s'appliquent à eux. Dans la plupart des domaines de compétences, l'échelle s'étend de 1 (très rarement) à 5 (très souvent). Dans le cas de la capacité d'argumenter et de la capacité d'adopter d'autres points de vue que le sien, l'échelle utilisée va de 1 (ne s'applique pas du tout) à 4 (s'applique tout à fait). Le tableau IV.11 montre les moyennes et les écarts-types obtenus dans les domaines de compétence qui ont passé avec succès le test de fiabilité statistique.

Tableau IV.11: Valeurs moyennes sur les échelles des compétences transversales

Domaines de compétences	Minimum	Maximum	Moyenne	Ecart-type
Planification du temps de travail	1,00	5,00	2,64	.961
Efficacité de la préparation des examens	1,00	5,00	3,66	.666
Gestion du temps	1,00	5,00	3,76	.646
Concentration	1,00	5,00	2,86	.700
Capacité d'analyse et de synthèse	1,00	5,00	3,29	.777
Capacité d'argumenter	1,00	4,00	3,28	.513
Capacité d'adopter d'autres points de vue que le sien	1,00	4,00	2,99	.578

N max = 13 358 (pondéré)

Les élèves ont jugé leurs propres compétences transversales de façon très différenciée. Ils ne planifient que «rarement» à «parfois» leur temps de travail (moyenne = 2,64). Il n'arrive également que «rarement» à «parfois» qu'ils parviennent à se concentrer sur une question (moyenne = 2,86). Mais ils n'éprouvent apparemment pas de grandes difficultés à préparer efficacement leurs examens (moyenne = 3,66) et plus généralement à bien gérer leur temps (moyenne = 3,76). Les élèves s'estiment aptes à travailler en équipe également. Ils se disent «plutôt» capables d'adopter le point de vue d'autres membres d'un groupe et, de façon encore plus marquée, d'argumenter (moyenne = 3,28).

Le tableau IV.12 montre comment les élèves ont répondu à diverses affirmations testant leur persévérance. Les résultats sont fournis séparément pour chaque item car il n'a pas été possible de les agréger dans une échelle statistiquement fiable. En moyenne, les élèves se jugent «souvent» persévérants (items 1 à 3) et ne croient que «rarement» qu'ils investissent plus de temps que la moyenne dans leurs études (item 4).

Tableau IV.12: Valeurs moyennes des questions sur la persévérance (1 = très rarement; 5 = très souvent)

Affirmations	Moyenne	Ecart-type
Je n'abandonne pas, même quand la matière est très difficile ou compliquée.	3,70	0,839
J'étudie aussi tard le soir ou durant le week-end quand c'est nécessaire.	3,85	1,086
Je passe le temps qu'il faut à étudier, jusqu'à ce que j'aie vraiment compris la matière.	3,63	0,836
Je consacre plus de temps à étudier que la plupart des autres élèves.	2,27	1,087

N max = 13 350 (pondéré)

Dans l'ensemble, les bacheliers et bacheliers s'estiment persévérants à l'école. Ils pensent que c'est le cas des autres élèves aussi. Ils disent en effet qu'ils ne passent pas davantage de temps que les autres à étudier (moyenne = 2,27). Il semble courant chez les bacheliers et bacheliers d'étudier tard le soir ou durant le week-end lorsque c'est indispensable (moyenne = 3,85). Dans leur grande majorité, ils affirment qu'ils ne lâchent pas prise facilement, même quand la matière est compliquée (moyenne 3,70) et qu'ils passent le temps qu'il faut pour vraiment comprendre une question (moyenne = 3,63).

5.7.3 Intérêt pour les différentes disciplines et valeur propédeutique du gymnase

En Suisse alémanique, les élèves ont également été interrogés sur leur intérêt pour les trois disciplines testées dans EVAMAR II (mathématiques, allemand et biologie). Ils devaient dire aussi s'ils estiment que l'école les prépare bien aux études supérieures. Ce volet de l'étude n'a pas été mené en Suisse romande et au Tessin par manque de temps.

Treize items ont servi à évaluer l'intérêt des élèves pour les différentes disciplines. Les analyses factorielles ont permis de distinguer diverses dimensions de la motivation: motivation intrinsèque, motivation instrumentale et orientation d'après la performance. Mais les 13 items peuvent aussi être agrégés dans une échelle unique: l'intérêt pour la discipline (coefficients d'homogénéité: entre 0,90 et 0,91). Les moyennes obtenues pour chacune des disciplines sont reproduites dans le tableau IV.13 (valeur élevée = intérêt élevé).

Tableau IV.13: Intérêt pour la matière enseignée

Discipline	Minimum	Maximum	Moyenne	Ecart-type
Mathématiques	1,00	4,00	2,04	.600
Biologie	1,00	4,00	2,41	.606
Langue première (allemand)	1,00	4,00	2,27	.633

N max = 9 293 (pondéré)

Globalement, la biologie est la discipline qui intéresse le plus les élèves. Les mathématiques les intéressent le moins. Ces résultats sont statistiquement significatifs. L'enquête qui avait été menée dans le cadre d'EVAMAR I avait débouché sur la même hiérarchie (voir Ramseier et al., 2006, p. 73).

Trois échelles s'adossant à l'étude TOSCA (voir Köller, Watermann, Trautwein & Lüdtke, 2004) ont été utilisées pour mesurer la valeur propédeutique des études gymnasiales. La première échelle (5 items) cherche à évaluer dans quelle mesure diverses activités courantes à l'université (récolter des informations, rédiger des textes à l'ordinateur ou recourir aux médias électroniques pour s'informer par exemple) sont déjà exercées au gymnase (réponses possibles: de 1 = jamais à 4 = souvent). Cette échelle a pu être reproduite avec un degré d'homogénéité satisfaisant (coefficient alpha de Cronbach = 0,72). Sa moyenne se monte à 2,86 (écart-type = 0,56), ce qui est supérieur à la moyenne théorique. En moyenne, les activités typiquement universitaires le sont donc «parfois» déjà au gymnase. La plus fréquente est de s'informer par le biais des médias électroniques selon

les gymnasiens (moyenne = 3,43, écart-type = 0,82). Il est plutôt rare en revanche qu'ils utilisent l'ordinateur pour produire des figures (moyenne = 2,23, écart-type = 0,87).

La deuxième échelle s'intéresse à d'autres activités qui préparent aux études. Elle n'a toutefois pas pu être reproduite. L'item qui obtient le score le plus élevé est la présentation d'exposés (moyenne = 3,64, écart-type = 0,57). Il est plutôt rare en revanche que les élèves préparent des expériences scientifiques (moyenne = 1,91, écart-type = 0,89) ou qu'ils rédigent des procès-verbaux (moyenne = 1,92, écart-type = 0,92).

Enfin, les élèves ont donné leur avis sur l'utilité de leur passage au gymnase pour les préparer aux exigences du monde universitaire. Ils ont été confrontés pour cela à des affirmations comme «J'ai compris au gymnase comment on aborde un problème scientifiquement» ou «J'ai bien développé ma culture générale au gymnase» (réponses possibles: de 1 = pas du tout à 4 = tout à fait). L'homogénéité de l'échelle n'est pas tout à fait satisfaisante, mais encore acceptable (coefficient de 0,68). La moyenne s'établit à 2,89 (écart-type = 0,43). Le jugement d'ensemble semble donc plutôt positif. Cela s'applique en particulier à l'item «J'ai le sentiment que le gymnase m'a bien préparé pour des études supérieures» (moyenne = 2,78, écart-type = 0,73). Les élèves sont le plus critique sur l'exactitude de l'affirmation suivante: «Le contenu des cours au gymnase m'a aidé à décider dans quelles études supérieures je veux me lancer». La moyenne sur cet item ne dépasse pas 2,34, mais l'écart-type est relativement élevé (0,98).

5.8 Interprétation des résultats sous l'angle de l'aptitude aux études

Les bacheliers et bacheliers suisses ont obtenu des résultats que l'on peut qualifier de normaux pour des tests standardisés qui ont été développés pour une population spécifique. Environ la moitié des items ont été résolus correctement. Le niveau de formation des élèves peut donc être qualifié de satisfaisant dans les gymnases suisses. Les résultats sont légèrement meilleurs en langue première qu'en mathématiques et en biologie. En biologie, les élèves qui avaient terminé leurs cours depuis au moins six mois ont été nettement moins performants au test, probablement sous l'effet de l'oubli qui s'instaure au fil du temps. Les connaissances emmagasinées au gymnase ne semblent acquises dans leur intégralité que pendant une courte période. Il est donc souhaitable que la fin des cours au gymnase intervienne aussi tard que possible avant les examens de maturité et que les élèves entament rapidement leurs études supérieures dans la foulée. Les résultats du test de compétences transversales (TCT) sont légèrement moins bons que les résultats du test qui est réalisé chaque année pour mesurer l'aptitude des candidats aux études de médecine. Cet écart est plausible dans la mesure où des élèves qui estiment ne pas avoir les capacités pour entamer des études de médecine ont aussi participé aux tests d'EVAMAR. La faiblesse de l'écart est toutefois un indice fort que les bacheliers et bacheliers testés se sont donné beaucoup de peine pour répondre aux questions qui leur étaient soumises. Dans le cas contraire, il serait difficilement explicable que les résultats du TCT soient si proches de ceux obtenus dans le cadre d'un test qui conditionne aussi fortement l'avenir personnel que celui d'aptitude aux études de médecine.

Les résultats des tests varient de façon considérable entre les élèves et entre les classes. Les écarts sont plus importants qu'aux examens de maturité. Une part substantielle des bacheliers et des bacheliers a donc vraisemblablement des compétences insuffisantes dans au moins un des domaines testés. L'analyse des résultats des examens de maturité – où les écoles attribuent les notes elles-mêmes – confirme ce point de vue, surtout en mathématiques. En 2007, 41,4% des élèves ont obtenu une note insuffisante lors de l'examen écrit dans cette discipline. La note finale de mathématiques, qui tient aussi compte des notes obtenues durant l'année et des résultats de l'examen oral, reste insuffisante dans 24,4% des cas. La question de savoir si les résultats insuffisants enregistrés lors des examens de maturité trahissent des déficits de connaissances et de compétences importants sous l'angle de l'aptitude générale aux études reste posée. Dans des cas extrêmes, des bacheliers et bacheliers ayant des notes insuffisantes disposent peut-être quand même d'aptitudes suffisantes dans les domaines qui comptent; à l'inverse, certains élèves qui ont obtenu des notes de maturité suffisantes ne sont peut-être pas dotés de toutes les compétences nécessaires dans les domaines cruciaux du point de vue de leur aptitude générale aux études. En revanche, les résultats en partie médiocres obtenus aux tests trahissent

clairement des déficits déterminants pour l'aptitude aux études dans au moins une discipline. Il n'est pas possible d'établir de façon déterministe à partir de quel niveau de performance aux tests des études supérieures couronnées de succès deviendraient illusoires, à moins de mener des enquêtes empiriques étendues et de longue haleine. Mais en tenant compte aussi des notes obtenues aux examens de maturité, on peut raisonnablement penser qu'une partie des bacheliers et bacheliers ne disposent pas de compétences suffisantes pour pouvoir se lancer dans des études supérieures dans n'importe quelle discipline – à moins qu'ils ne comblient leurs lacunes par ailleurs. Pourtant, le certificat de maturité est censé attester leur aptitude générale aux études.

Les gymnases ne produisent donc pas que des bacheliers et bacheliers dont les compétences générales peuvent être qualifiées d'au moins suffisantes. On pouvait déjà s'en douter d'après le taux de notes insuffisantes délivrées dans les gymnases aux examens de maturité en mathématiques et à l'écrit de langue première. Malgré un bon niveau de formation, des améliorations sont donc a priori possibles. Cette conclusion n'a rien d'extraordinaire. Toute évaluation des performances à l'aide de tests standardisés conclura nécessairement que des améliorations sont possibles, quelle que soit le niveau de scolarité étudié. Des mesures fortes se seraient imposées si nettement moins de la moitié des items avaient été résolus correctement et si l'estimation des aptitudes personnelles moyennes avait été largement inférieure à un niveau moyen. Ce n'est pas le cas.

Cela dit, la maturité est censée attester que son titulaire dispose du bagage nécessaire pour entamer n'importe quelles études universitaires. Or ce n'est sans doute pas le cas pour une frange non négligeable des bacheliers et des bacheliers. L'évaluation EVAMAR II le met en évidence dans les disciplines qu'elle a testées, mais le tableau d'ensemble ne serait sans doute guère différent en étendant les recherches à d'autres domaines. Cette conclusion n'est pas nouvelle non plus. Les lacunes existantes, favorisées par le système de compensation des notes du RRM, s'expriment déjà dans les notes de maturité. Cela permet d'expliquer au moins en partie que le corps professoral se plaigne de lacunes chez leurs nouveaux étudiantes et étudiants, surtout en langue première et en mathématiques (voir chapitre II, section 3.3) et que les taux d'échecs lors des premiers examens universitaires soient parfois considérables.

Les résultats du questionnaire complémentaire montrent que les élèves se font une image incomplète de leur capacité d'apprendre et de travailler. Cela confirme les résultats d'autres évaluations et ceux de l'enquête réalisée auprès du corps professoral (voir sous-projet A4, chapitre II).

Pour pouvoir envisager que tous les bacheliers et bacheliers soient un jour «parfaitement aptes aux études», il faudrait certainement changer les règles du jeu aux examens de maturité; les notes insuffisantes ne devraient plus être tolérées. Mais cela aurait vraisemblablement deux conséquences indésirables: l'abaissement du taux de maturités et des changements dans la pratique d'évaluation. Selon toute vraisemblance, une «aptitude parfaite aux études» de toutes les bacheliers et de tous les bacheliers restera donc une fiction.

6 Analyses comparatives selon plusieurs critères

6.1 Généralités

Cette section compare les aptitudes personnelles des élèves (ou dans différents gymnases) d'après leurs résultats aux différents tests d'EVAMAR II à l'aune de différents critères successifs. Cet exercice revêt un intérêt particulier dans un pays où le système gymnasial est aussi hétérogène qu'en Suisse.

Les comparaisons portent d'abord sur les différentes strates géographiques (point 6.2). Les élèves sont ensuite regroupés en sous-populations sur la base d'autres critères, comme l'option spécifique choisie ou le sexe (à partir du point 6.3).

Les comparaisons ont été faites selon la méthode de calcul P, autrement dit, d'après l'estimation des aptitudes personnelles selon le modèle de Rasch (voir point 5.1). Les données ont été standardisées pour faciliter les comparaisons entre les tests. La moyenne pour l'ensemble de la Suisse a été fixée à 500 points. Environ deux tiers des bacheliers et bachelères (68,2%) ont un niveau d'aptitudes personnelles qui les situent dans la fourchette entre 400 et 600 points. L'intervalle entre 300 et 700 points englobe 95,4% des élèves, et seuls 2,3% d'entre eux obtiennent moins de 200 ou plus de 800 points. Pratiquement aucun élève ne se situe en dessous de 200 points ou au-delà de 800.

Les différences entre les sous-populations ont été systématiquement soumises à un test de signification. Le test *t* a été appliqué lorsqu'il n'y avait que deux sous-populations à comparer et le test a posteriori de Duncan lorsqu'il y en avait davantage (comparaisons multiples de moyennes). Des groupes de sous-populations significativement différents ont ainsi pu être identifiés (niveau de 5% au moins, $\alpha = 0,05$). Chacun de ces groupes s'est vu attribuer un numéro, en partant du groupe avec les aptitudes les plus basses (groupe 1). Certaines sous-populations sont rattachées à deux groupes différents, ce qui signifie qu'elles ne sont significativement différentes des sous-populations d'aucun des ces deux groupes.

Des différences de moyenne même ténues et guère pertinentes en pratique peuvent néanmoins être significatives en raison de la grande taille des échantillons. C'est pourquoi nous avons calculé aussi pour chaque test l'amplitude de l'effet selon Cohen (voir Bortz & Döring, 2006). L'amplitude de l'effet correspond au rapport entre la différence entre les moyennes et l'écart-type moyen de deux entités comparées (valeur *d*). L'amplitude de l'effet qui nous intéressera dans ce qui suit est l'amplitude maximale, soit celle entre les deux sous-populations dont les moyennes sont les plus éloignées l'une de l'autre (*d* max). La valeur *d* est généralement interprétée de la manière suivante (voir Bortz & Döring, 2006):

- $d = 0,20$: petit effet
- $d = 0,50$: effet moyen
- $d = 0,80$: grand effet

On ne peut en aucun cas déduire que la qualité de l'enseignement est plus mauvaise dans les gymnases où les moyennes sont basses. Les causes sont en général structurelles!

6.2 Comparaisons selon la strate géographique

6.2.1 Résultats des disciplines testées

6.2.1.1 Langue première

Le niveau des aptitudes personnelles dans les différentes strates, estimé sur la base des résultats du test de langue première, figure dans les tableaux IV.14 à IV.18.

Tableau IV.14: Comparaison des résultats au test de langue première selon la strate géographique (total)

Strates	N (pondéré)	Groupes significatifs (1,2,3)		
		1	2	3
Suisse alémanique, gym3	1844	489		
ZH	1921		498	
Suisse romande 1 (gym3)	3187		501	
Suisse alémanique, grands	4442		501	
Suisse alémanique, petits	923		502	
Suisse romande 2	1041			524

d max = 0,41

Tableau IV.15: Comparaison des résultats au test de langue première selon la strate géographique
(compréhension générale de l'écrit)

Strates	N (pondéré)	Groupes significatifs (1,2,3)		
		1	2	3
Suisse alémanique, gym3	1844	495		
Suisse alémanique, grands	4442	498		
Suisse romande 1 (gym3)	3187	499		
ZH	1921	500		
Suisse alémanique, petits	923		507	
Suisse romande 2	1041			514

d max = 0,24

Tableau IV.16: Comparaison des résultats au test de langue première selon la strate géographique
(compréhension détaillée de l'écrit)

Strates	N (pondéré)	Groupes significatifs (1,2,3)				
		1	2	3	4	5
Suisse alémanique, gym3	1844	487				
Suisse romande 1 (gym3)	3187		494			
ZH	1921			502		
Suisse alémanique, grands	4442			504	504	
Suisse alémanique, petits	923				510	
Suisse romande 2	1041					519

d max = 0,38

Tableau IV.17: Comparaison des résultats au test de langue première selon la strate géographique
(fonctionnement de la langue/vocabulaire)

Strates	N (pondéré)	Groupes significatifs (1,2,3)		
		1	2	3
Suisse alémanique, gym3	1844	492		
Suisse alémanique, grands	4442	493		
ZH	1921		499	
Suisse alémanique, petits	923		501	
Romandie 1 (MD3)	3187			512
Romandie 2	1041			517

d max = 0,34

Tableau IV.18: Comparaison des résultats au test de langue première selon la strate géographique
(fonctionnement de la langue/compétence grammaticale et orthographe)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
Suisse alémanique, gym3	1844	489		
Suisse alémanique, grands	4442		500	
Suisse romande 1 (gym3)	3187		500	
ZH	1921		502	
Suisse alémanique, petits	923			509
Suisse romande 2	1041			512

d max = 0,32

Les établissements alémaniques où le gymnase dure trois ans ont obtenu les moins bons résultats en langue première (voir tableau IV.14). Ce résultat est significatif. L'amplitude maximale de l'effet est petite à moyenne par rapport à la meilleure strate («Suisse romande 2»). Des distinctions plus fines sont possibles dans les différents domaines de compétences:

- Les petits cantons alémaniques et «Suisse romande 2» sont les deux seules strates à sortir du lot en compréhension générale de l'écrit.
- Un palier de différenciation supplémentaire apparaît en compréhension détaillée de l'écrit. «Suisse romande 1» forme un groupe en soi en queue de classement.
- En vocabulaire, les grands cantons alémaniques tombent à leur tour dans le groupe des strates les plus mal classées.
- Les petits cantons alémaniques rejoignent les meilleurs en grammaire et en orthographe. Pour le reste, la répartition des strates dans ce domaine de compétences est conforme au tableau d'ensemble.

Tendanciellement, les résultats sont plutôt plus mauvais dans les gymnases où la durée minimale des études n'est que de trois ans, du moins au sein de chaque région linguistique.

6.2.1.2 Mathématiques

Le niveau des aptitudes personnelles dans les différentes strates, estimé sur la base des résultats du test de mathématiques, figure dans les tableaux IV.14 à IV.18.

Tableau IV.19: Comparaison des résultats au test de mathématiques selon la strate géographique (total)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Suisse romande 1 (gym3)	3180	477				
Suisse alémanique, gym3	1865		492			
Suisse alémanique, grands	4446		496	496		
ZH	2050			501	501	
Suisse alémanique, petits	932				505	
Suisse romande 2	1048					522

d max = 0,48

Tableau IV.20: Comparaison des résultats au test de mathématiques selon la strate géographique (calcul analytique)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1			1	
Suisse romande 1 (gym3)	3180	483				
Suisse alémanique, gym3	1865		492			
Suisse alémanique, grands	4446		496	496		
ZH	2050			499		
Suisse alémanique, petits	932				505	
Suisse romande 2	1048					520

d max = 0,41

Tableau IV.21: Comparaison des résultats au test de mathématiques selon la strate géographique (géométrie)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 1 (gym3)	3180	488			
Suisse alémanique, grands	4446		494		
Suisse alémanique, gym3	1865		498	498	
ZH	2050			501	
Suisse alémanique, petits	932			503	
Suisse romande 2	1048				514

d max = 0,28

Tableau IV.22: Comparaison des résultats au test de mathématiques selon la strate géographique (stochastique)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 1 (gym3)	3180	482			
Suisse alémanique, gym3	1865		492		
ZH	2050			498	
Suisse alémanique, petits	932			501	
Suisse alémanique, grands	4446			501	
Suisse romande 2	1048				528

d max = 0,55

Tableau IV.23: Comparaison des résultats au test de mathématiques selon la strate géographique (divers)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Suisse romande 1 (gym3)	3180	477				
Suisse alémanique, gym3	1865		489			
Suisse alémanique, petits	932			498		
Suisse alémanique, grands	4446			500	500	
ZH	2050				505	
Suisse romande 2	1048					517

d max = 0,47

En mathématiques, la durée du gymnase semble avoir un impact encore plus important sur le niveau des performances qu'en langue première. Les établissements où la durée minimale des études n'est que de trois ans s'en sortent nettement moins bien que ceux où elle est d'au moins quatre ans, spécialement en Suisse romande. L'amplitude de l'effet est globalement moyenne entre la strate avec les résultats plus bas («Suisse romande 1») et celle où ils sont les plus élevés («Suisse romande 2»). En stochastique et dans les domaines regroupés sous «Divers», l'amplitude de l'effet est légèrement plus accentuée que la moyenne, et inversement en géométrie.

6.2.1.3 Biologie

Le niveau des aptitudes personnelles dans les différentes strates, estimé sur la base des résultats du test de biologie, figure dans les tableaux IV.24 à IV.30.

Tableau IV.24: Comparaison des résultats au test de biologie selon la strate géographique (total)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 1 (gym3)	3152	459			
Suisse romande 2	1048	462			
Suisse alémanique, gym3	1884		506		
Suisse alémanique, petits	916		510	510	
Suisse alémanique, grands	4464			514	
ZH	1974				522

d max = 0,75

Tableau IV.25: Comparaison des résultats au test de biologie selon la strate géographique (évolution)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 1 (gym3)	3152	462			
Suisse romande 2	1048	464			
Suisse alémanique, gym3	1884		506		
Suisse alémanique, petits	916			512	
Suisse alémanique, grands	4464			514	514
ZH	1974				519

d max = 0,74

Tableau IV.26: Comparaison des résultats au test de biologie selon la strate géographique (structure et fonction)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 1 (gym3)	3152	463			
Suisse romande 2	1048	463			
Suisse alémanique, gym3	1884		505		
Suisse alémanique, petits	916		509		
Suisse alémanique, grands	4464			515	
ZH	1974				522

d max = 0,74

Tableau IV.27: Comparaison des résultats au test de biologie selon la strate géographique (génétique)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 2	1048	463			
Suisse romande 1 (gym3)	3152	464			
Suisse alémanique, gym3	1884		505		
Suisse alémanique, petits	916		509	509	
Suisse alémanique, grands	4464			513	
ZH	1974				521

d max = 0,76

Tableau IV.28: Comparaison des résultats au test de biologie selon la strate géographique (traitement de l'information)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 1 (gym3)	3152	462			
Suisse romande 2	1048	463			
Suisse alémanique, gym3	1884		504		
Suisse alémanique, petits	916			509	
Suisse alémanique, grands	4464			513	
ZH	1974				522

d max = 0,75

Tableau IV.29: Comparaison des résultats au test de biologie selon la strate géographique (écologie)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 1 (gym3)	3152	460			
Suisse romande 2	1048	463			
Suisse alémanique, gym3	1884		506		
Suisse alémanique, petits	916		511	511	
Suisse alémanique, grands	4464			513	
ZH	1974				521

d max = 0,75

Tableau IV.30: Comparaison des résultats au test de biologie selon la strate géographique (métabolisme)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Suisse romande 1 (gym3)	3152	461			
Suisse romande 2	1048	462			
Suisse alémanique, gym3	1884		505		
Suisse alémanique, petits	916		510	510	
Suisse alémanique, grands	4464			514	
ZH	1974				522

d max = 0,75

En biologie, le canton de ZH est clairement en tête. L'amplitude maximale de l'effet est marquée. Les écoles de Suisse romande sont groupées en queue de classement. La durée du gymnase y joue un rôle insignifiant. En Suisse alémanique en revanche, les gymnases où les études durent au moins quatre ans s'en sortent une nouvelle fois mieux dans la règle, que ceux où cette durée est de trois ans.

Si l'on isole les établissements où les cours de biologie avaient pris fin depuis des mois au moment du test, la hiérarchie reste à peu près la même (voir tableau IV.31). L'amplitude maximale de l'effet tend même à se renforcer. L'effet d'oubli des connaissances acquises accentue probablement les différences. Dans les gymnases où les cours de biologie n'étaient pas encore terminés au moment du test, les différences entre les strates sont beaucoup plus faibles mais le niveau d'ensemble est sensiblement plus élevé (tableau IV.32). Dans la plupart des cas, les élèves concernés sont des bacheliers et bacheliers qui ont choisi la biologie en option (spécifique ou complémentaire). Chez ces élèves aussi, un certain écart subsiste entre les résultats enregistrés dans l'ouest et dans l'est du pays.

Tableau IV.31: Comparaison des résultats au test de biologie des élèves qui avaient déjà terminé les cours dans cette discipline selon la strate géographique (total)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)				
		1	2	3	4	5
Suisse romande 1 (gym3)	2293	436				
Suisse romande 2	840		447			
Suisse alémanique, gym3	952			493		
Suisse alémanique, petits	718			495		
Suisse alémanique, grands	3409				502	
ZH	1478					511

d max = 0,97

Tableau IV.32: Comparaison des résultats au test de biologie des élèves qui n'avaient pas encore terminé les cours dans cette discipline selon la strate géographique (total)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)	
		1	2
Suisse alémanique, gym3	932	519	
Suisse romande 2	207	520	
Suisse romande 1 (gym3)	858	522	
ZH	495		554
Suisse alémanique, grands	1054		554
Suisse alémanique, petits	198		565

d max = 0,56

6.2.2 Notes de maturité

Tableau IV.33: Comparaison des notes de maturité en langue première selon la strate géographique (note finale)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)				
		1	2	3	4	5
Suisse romande 1 (gym3)	3151	4,48				
Suisse romande 2	1043		4,55			
Suisse alémanique, gym3	1628			4,63		
Suisse alémanique, grands	4514				4,67	
ZH	2091					4,73
Suisse alémanique, petits	959					4,76

d max = 0,50

En langue première, la note finale de maturité varie significativement d'une strate à l'autre également. L'amplitude maximale de l'effet est moyennement prononcée. La hiérarchie n'est pas la même que pour les résultats du test. Il n'y a pas d'explication plausible évidente à ce phénomène. Le fait que les connaissances et les compétences enseignées en langue première soient beaucoup plus vastes et en partie différentes de celles testées dans le cadre d'EVAMAR II joue peut-être un rôle. Il est frappant de constater que les notes délivrées en langue première sont les plus basses en Suisse occidentale, et ce de part et d'autre de la frontière des langues (Suisse romande et canton de Berne).

Tableau IV.34: Comparaison des notes de maturité en mathématiques selon la strate géographique (note finale)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)	
		1	2
Suisse romande 1 (gym3)	3151	4,29	
Suisse alémanique, gym3	1628	4,30	
Suisse alémanique, grands	4514	4,35	4,35
Suisse romande 2	1043		4,38
Suisse alémanique, petits	959		4,39
ZH	2091		4,39

d max = 0,12

L'écart entre la strate avec la note finale de mathématiques moyenne la plus basse et celle où elle est la plus élevée est significatif, bien qu'il se limite à un dixième de point en chiffres absolus. L'amplitude de l'effet est très faible.

Tableau IV.35: Comparaison des notes de maturité en biologie selon la strate géographique (notation continue)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)			
		1	2	3	4
ZH	2091	4,58			
Suisse alémanique, petits	826		4,63		
Suisse alémanique, gym3	1471		4,64		
Suisse romande 2	1043			4,70	
Suisse alémanique, grands	4507			4,73	
Suisse romande 1 (gym3)	1172				4,77

d max = 0,36

Tableau IV.36: Comparaison des notes de maturité en sciences expérimentales selon la strate géographique (note finale)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
Suisse romande 1 (gym3)	3151	4,52		
ZH	2067	4,54		
Suisse alémanique, gym3	1628		4,60	
Suisse alémanique, petits	964		4,61	
Suisse romande 2	1043		4,63	4,63
Suisse alémanique, grands	4514			4,66

d max = 0,26

En biologie, la hiérarchie entre les strates qui résulte de la notation continue est grosso modo l'inverse de celle qui se dégage des résultats du test. L'amplitude de l'effet est d'un niveau (presque) moyen. La note finale de sciences expérimentales ne diffère que peu d'une strate à l'autre (voir tableau IV.36).

Dans l'ensemble, les différences entre les strates sont faibles voire inexistantes sur les examens de maturité. Cela s'explique probablement par le fait qu'il n'existe pas – et ne peut pas exister – de pratique d'évaluation unifiée en Suisse.

6.2.3 Résultats du test des compétences transversales (TCT)

Les cinq prochains tableaux (IV.37 à IV.40) comparent les compétences transversales qui ont été évaluées dans le cadre du TCT entre les strates. Les élèves de Suisse romande n'ont pas été soumis à ce test (voir point 1.1).

Tableau IV.37: Comparaison des résultats du test des compétences transversales selon la strate géographique (TCT, quantités et formes)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
Suisse alémanique, gym3	1821	487		
Suisse alémanique, grands	4471		496	
Suisse alémanique, petits	933			502
ZH	2053			506

d max = 0,23

Tableau IV.38: Comparaison des résultats du test des compétences transversales selon la strate géographique (TCT, diagrammes et tableaux)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)			
		1	2	3	4
Suisse alémanique, gym3	1821	487			
Suisse alémanique, grands	4471		495		
Suisse alémanique, petits	933			501	
ZH	2053				507

d max = 0,25

Tableau IV.39: Comparaison des résultats du test des compétences transversales selon la strate géographique (TCT, compréhension de textes)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
Suisse alémanique, gym3	1821	486		
Suisse alémanique, grands	4471		496	
Suisse alémanique, petits	933		501	
ZH	2053			507

d max = 0,24

Tableau IV.40: Comparaison des résultats du test des compétences transversales selon la strate géographique (TCT, notions élémentaires de sciences expérimentales)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
Suisse alémanique, gym3	1821	486		
Suisse alémanique, grands	4471		496	
Suisse alémanique, petits	933		500	
ZH	2053			507

d max = 0,25

La durée du gymnase n'influence pas seulement les résultats des tests dans les différentes disciplines mais aussi ceux du test des compétences transversales. Les gymnases où la maturité peut être décrochée après trois ans d'études déjà ont enregistré des résultats significativement plus mauvais que les autres. Dans le cas du TCT, l'amplitude de l'effet est toutefois faible. Les aptitudes qu'il évalue semblent relativement indépendantes de connaissances spécifiques liées à certaines disciplines. Cela confirme que le TCT a bien une vocation interdisciplinaire.

6.2.4 Résultats choisis issus du questionnaire complémentaire

Les tableaux IV.41 à IV.43 montrent la façon dont les élèves des différentes strates ont évalué leurs propres compétences transversales dans le questionnaire complémentaire. Seules les compétences pour lesquelles la différence entre les moyennes extrêmes est au moins de faible amplitude sont listées ($d > 0.2$).

Tableau IV.41: Comparaison des résultats de l'auto-évaluation des compétences transversales selon la strate géographique (planification du temps de travail: 1 = mauvaise, 5 = très bonne)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
Suisse romande 1 (gym3)	3153	2,55		
ZH	2022	2,56		
Suisse alémanique, gym3	1879		2,67	
Suisse alémanique, petits	916		2,68	
Suisse alémanique, grands	4388		2,68	
Suisse romande 2	1035			2,78

d max = 0,24

Tableau IV.42: Comparaison des résultats de l'auto-évaluation des compétences transversales selon la strate géographique (efficacité de la préparation aux examens: 1 = préparation inefficace, 5 = préparation très efficace)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)			
		1	2	3	4
ZH	2022	3,56			
Suisse alémanique, gym3	1879	3,58			
Suisse alémanique, grands	4388		3,64		
Suisse alémanique, petits	916		3,68		
Suisse romande 1 (gym3)	3153			3,75	
Suisse romande 2	1035				3,84

d max = 0,41

Tableau IV.43: Comparaison des résultats de l'auto-évaluation des compétences transversales selon la strate géographique (gestion du temps: 1 = très mauvaise, 5 = très bonne)

Strates	N (pondéré)	Groupes significatifs (1, 2, 3)				
		1	2	3	4	5
Suisse romande 1 (gym3)	3153	3,62				
Suisse romande 2	1035		3,70			
Suisse alémanique, gym3	1879			3,77		
ZH	2022			3,79	3,79	
Suisse alémanique, grands	4388				3,83	
Suisse alémanique, petits	916					3,90

d max = 0,43

Les différences entre les strates ne sont prêtes guère à être interprétées. Les écarts sont tout au plus d'amplitude moyenne. Les moyennes sont très voisines (elles se situent toutes entre 3,5 et 4). Aucune hiérarchie solide ne se dégage entre les strates.

6.3 Comparaisons selon l'option spécifique choisie

L'un des objectifs principaux de la réforme de la maturité en 1995 était d'offrir aux bacheliers et bacheliers un choix plus large d'options spécifiques (voir chapitre I, section 1). Constitué des disciplines fondamentales, le tronc commun a pour vocation de consolider la culture générale des élèves. Mais les gymnasiennes et gymnasiens devaient pouvoir approfondir leur formation gymnasiale de façon plus individualisée, en choisissant une option spécifique – et plus tard une option complémentaire – en adéquation avec leurs centres d'intérêts. Indépendamment des options choisies, les bacheliers et bacheliers sont censés développer une aptitude générale aux études, donc d'être capables de se lancer dans des études supérieures dans n'importe quelle branche. L'étude EVAMAR I (Ramseier et al., 2004, p. 138) et d'autres enquêtes de terrain (Oelkers, 2008, p. 120) ont montré que les connaissances et les capacités des élèves dans les disciplines fondamentales ne se distinguent guère au gymnase. Cependant, certaines options spécifiques canalisent les élèves vers certaines branches à l'université. Dans ce contexte, il est du plus grand intérêt de voir si ces conclusions se reflètent dans les performances des élèves mesurées dans le cadre d'EVAMAR II.

6.3.1 Résultats des disciplines testées

6.3.1.1 Langue première

Le niveau des aptitudes personnelles dans les différentes sous-populations, estimé sur la base des résultats du test de langue première, figure dans les tableaux IV.44 à IV.48.

Tableau IV.44: Comparaison des résultats au test de langue première selon l'option spécifique choisie (total)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Arts visuels	1176	477				
Economie et droit	3097		492			
Musique	762		494			
Physique et applications des mathématiques	1403			501		
Langues modernes	3217			502		
Biologie et chimie	1722				510	
Philosophie/Pédagogie/Psychologie	1312				514	
Langues anciennes	669					533

d max = 0,59

Tableau IV.45: Comparaison des résultats au test de langue première selon l'option spécifique choisie (compréhension générale de l'écrit)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Arts visuels	1176	484			
Economie et droit	3097	489			
Musique	762		495		
Langues modernes	3217		501		
Physique et applications des mathématiques	1403		501		
Philosophie/Pédagogie/Psychologie	1312			508	
Biologie et chimie	1722			510	
Langues anciennes	669				536

d max = 0,63

Tableau IV.46: Comparaison des résultats au test de langue première selon l'option spécifique choisie (compréhension détaillée de l'écrit)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5, 6)					
		1	2	3	4	5	6
Arts visuels	1176	485					
Economie et droit	3097		492				
Musique	762		493				
Physique et applications des mathématiques	1403		498	498			
Biologie et chimie	1722			503	503		
Langues modernes	3217				507	507	
Philosophie/Pédagogie/Psychologie	1312					511	
Langues anciennes	669						526

d max = 0,46

Tableau IV.47: Comparaison des résultats au test de langue première selon l'option spécifique choisie (fonctionnement de la langue/vocabulaire)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Musique	762	484				
Arts visuels	1176	485				
Economie et droit	3097		491			
Langues modernes	3217			503		
Physique et applications des mathématiques	1403			506	506	
Philosophie/Pédagogie/Psychologie	1312				510	
Biologie et chimie	1722				511	
Langues anciennes	669					527

d max = 0,58

Tableau IV.48: Comparaison des résultats au test de langue première selon l'option spécifique choisie (fonctionnement de la langue/compétence grammaticale et orthographe)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Arts visuels	1176	487				
Musique	762	488				
Economie et droit	3097	491	491			
Biologie et chimie	1722		494			
Physique et applications des mathématiques	1403			501		
Langues modernes	3217				509	
Philosophie/Pédagogie/Psychologie	1312				511	
Langues anciennes	669					531

d max = 0,59

Les élèves qui ont choisi l'option spécifique «Langues anciennes» (en général le latin) ont obtenu les meilleurs résultats dans tous les domaines en moyenne. A l'autre extrême, on trouve les bacheliers et bacheliers qui ont choisi les arts visuels. Ils font partie du groupe de queue dans tous les domaines, parfois en compagnie de la sous-population «Musique» (vocabulaire, compétence grammaticale et orthographe) et de la sous-population «Economie et droit» (compréhension générale de l'écrit, compétence grammaticale et orthographe). Les amplitudes maximales dans les différents sous-tests sont d'un niveau moyen. Un résultat frappant est le classement moyen des bacheliers et bacheliers qui ont choisi «Langues modernes» en option spécifique alors que le test portait sur la langue première.

6.3.1.2 Mathématiques

Le niveau des aptitudes personnelles dans les différentes sous-populations, estimé sur la base des résultats du test de mathématiques, figure dans les tableaux IV.49 à IV.53. Certaines bacheliers et certains bacheliers, ceux qui ont choisi l'option spécifique «Physique et application des mathématiques», ont reçu une formation avancée en mathématiques. Les amplitudes maximales ont donc été calculées de deux manières: en tenant compte de ces élèves (donnée principale) et sans en tenir compte (donnée indiquée entre parenthèse).

Tableau IV.49: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (total)

Options spécifiques	N (pondéré)	Groupes significatifs (1,2, 3, 4, 5, 6, 7)						
		1	2	3	4	5	6	7
Arts visuels	1156	455						
Philosophie/Pédagogie/Psychologie	1313		462					
Musique	698		467					
Langues modernes	3336			475				
Economie et droit	3181				485			
Biologie et chimie	1741					504		
Langues anciennes	671						512	
Physique et applications des mathématiques	1416							614

d max = 1,96 (0,69)

Tableau IV.50: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (calcul analytique)

Options spécifiques	N (pondéré)	Groupes significatifs (1,2, 3, 4, 5, 6)					
		1	2	3	4	5	6
Philosophie/Pédagogie/Psychologie	1313	462					
Arts visuels	1156	463					
Musique	698	465					
Langues modernes	3336		478				
Economie et droit	3181			490			
Biologie et chimie	1741				500		
Langues anciennes	671					509	
Physique et applications des mathématiques	1416						608

d max = 1,85 (0,57)

Tableau IV.51: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (géométrie)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Arts visuels	1156	462				
Philosophie/Pédagogie/Psychologie	1313	464				
Musique	698		478			
Langues modernes	3336		478			
Economie et droit	3181			488		
Biologie et chimie	1741				508	
Langues anciennes	671				508	
Physique et applications des mathématiques	1416					608

d max = 0,92 (0,59)

Tableau IV.52: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (stochastique)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5, 6, 7)						
		1	2	3	4	5	6	7
Arts visuels	1156	463						
Musique	698		472					
Philosophie/Pédagogie/Psychologie	1313		476					
Langues modernes	3336			484				
Economie et droit	3181				491			
Biologie et chimie	1741					501		
Langues anciennes	671						521	
Physique et applications des mathématiques	1416							582

d max = 1,43 (0,71)

Tableau IV.53: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (divers)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5, 6)					
		1	2	3	4	5	6
Arts visuels	1156	468					
Philosophie/Pédagogie/Psychologie	1313	468					
Musique	698		479				
Langues modernes	3336		481				
Economie et droit	3181			490			
Biologie et chimie	1741				498		
Langues anciennes	671					510	
Physique et applications des mathématiques	1416						586

d max = 1,54 (0,55)

Les bacheliers et bacheliers de l'option spécifique «Physique et applications des mathématiques» ont obtenu clairement les meilleurs résultats dans tous les domaines du test de mathématiques. Ce résultat était attendu, mais l'amplitude maximale de l'effet est chaque fois extrêmement importante. La sous-population «Arts visuels» est la moins performante dans tous les domaines, parfois en compagnie de la sous-population «Philosophie/Pédagogie/Psychologie» (calcul analytique, géométrie, divers) et de la sous-population «Musique» (calcul analytique). Les bacheliers et bacheliers de la sous-population «Langues anciennes» ont obtenu les meilleurs résultats après ceux de la sous-population «Physique et applications des mathématiques». Ils se classent devant ceux qui ont choisi l'option spécifique «Biologie et chimie» et l'option «Economie et droit». L'amplitude de l'effet entre les meilleurs «non spécialistes» («Langues anciennes») et la sous-population la plus faible est considérable. Elle varie entre 0,55 (domaine «Divers») et 0,71 (stochastique).

6.3.1.3 Biologie

Le niveau des aptitudes personnelles dans les différentes sous-populations, estimé sur la base des résultats de biologie, figure dans les tableaux IV.54 à IV.60. Certaines bacheliers et certains bacheliers, ceux qui ont choisi l'option spécifique «Biologie et chimie», ont reçu un enseignement spécifique plus poussé. Ils devraient disposer de connaissances et de capacités bien meilleures que les autres en biologie. C'est pourquoi l'amplitude maximale de l'effet est de nouveau indiquée de deux manières: en tenant compte de cette sous-population (donnée principale) et en en faisant abstraction (donnée entre parenthèses).

Tableau IV.54: Comparaison des résultats au test de biologie selon l'option spécifique choisie (total)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Musique	763	475				
Langues modernes	3260		485			
Philosophie/Pédagogie/Psychologie	1309		484			
Economie et droit	3137		487			
Arts visuels	1167		491	491		
Physique et applications des mathématiques	1402			496		
Langues anciennes	668				505	
Biologie et chimie	1723					559

d max = 1,08 (0,35)

Tableau IV.55: Comparaison des résultats au test de biologie selon l'option spécifique choisie (évolution)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Musique	763	483			
Economie et droit	3137	486			
Langues modernes	3260	486			
Philosophie/Pédagogie/Psychologie	1309	489			
Arts visuels	1167		496		
Physique et applications des mathématiques	1402		498		
Langues anciennes	668			504	
Biologie et chimie	1723				554

d max = 0,988 (0,27)

Tableau IV.56: Comparaison des résultats au test de biologie selon l'option spécifique choisie (structure et fonction)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Economie et droit	3137	484			
Musique	763	485			
Langues modernes	3260	485			
Philosophie/Pédagogie/Psychologie	1309	487			
Arts visuels	1167		497		
Physique et applications des mathématiques	1402		498		
Langues anciennes	668			506	
Biologie et chimie	1723				557

d max = 0,99 (0,27)

Tableau IV.57: Comparaison des résultats au test de biologie selon l'option spécifique choisie (génétique)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Economie et droit	3137	484			
Musique	763	484			
Langues modernes	3260	486			
Philosophie/Pédagogie/Psychologie	1309	487	487		
Arts visuels	1167		492		
Physique et applications des mathématiques	1402			499	
Langues anciennes	668			502	
Biologie et chimie	1723				557

d max = 0,99 (0,22)

Tableau IV.58: Comparaison des résultats au test de biologie selon l'option spécifique choisie (traitement de l'information)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5, 6)					
		1	2	3	4	5	6
Musique	763	481					
Langues modernes	3260	484	484				
Economie et droit	3137	484	484				
Philosophie/Pédagogie/Psychologie	1309		488	488			
Arts visuels	1167			492			
Physique et applications des mathématiques	1402				499		
Langues anciennes	668					506	
Biologie et chimie	1723						557

d max = 1,00 (0,30)

Tableau IV.59: Comparaison des résultats au test de biologie selon l'option spécifique choisie (écologie)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Musique	763	484				
Economie et droit	3137	484				
Langues modernes	3260	484				
Philosophie/Pédagogie/Psychologie	1309	488	488			
Arts visuels	1167		493	493		
Physique et applications des mathématiques	1402			498		
Langues anciennes	668				506	
Biologie et chimie	1723					555

d max = 0,95 (0,28)

Tableau IV.60: Comparaison des résultats au test de biologie selon l'option spécifique choisie (métabolisme)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Economie et droit	3137	483				
Musique	763	484				
Langues modernes	3260	485				
Philosophie/Pédagogie/Psychologie	1309	488	488			
Arts visuels	1167		492	492		
Physique et applications des mathématiques	1402			498		
Langues anciennes	668				505	
Biologie et chimie	1723					559

$d_{\max} = 1,02 (0,28)$

Comme on pouvait s'y attendre, les bacheliers et bachelères qui ont choisi l'option spécifique «Biologie et chimie» ont obtenu les meilleurs résultats dans tous les domaines. La sous-population «Langues anciennes» arrive de nouveau systématiquement en deuxième position. Elle est donc classée devant que la sous-population «Physique et applications mathématiques» qui est pourtant composée de «spécialistes» en sciences expérimentales. L'amplitude de l'effet est très marqué entre la sous-population «Biologie et chimie» et la sous-population la moins bonne. Il est modéré entre les meilleurs «non spécialistes» («Langues anciennes») et la sous-population la plus mauvaise: d varie entre 0,22 (génétique) et 0,35 (résultat global). La sous-population «Musique» est de loin la moins performante globalement. Dans les domaines particuliers, elle est systématiquement dans le groupe de queue, en compagnie des sous-populations «Langues modernes», et «Economie et droit». La sous-population «Philosophie/Pédagogie/Psychologie» fait aussi partie du groupe des plus mauvais dans tous les domaines, sauf dans le domaine «Traitement de l'information».

6.3.2 Notes de maturité

La comparaison des résultats des tests a montré qu'il existait de grandes différences entre les sous-populations. Les tableaux IV.61 à IV.64 montrent si ces écarts se vérifient lorsque l'on compare les notes de maturité dans les trois disciplines testées dans le cadre d'EVAMAR II.

Tableau IV.61: Comparaison des notes de maturité selon l'option spécifique choisie

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Economie et droit	3138	4,54				
Biologie et chimie	1713	4,58				
Arts visuels	1152	4,58				
Physique et applications des mathématiques	1421		4,63			
Langues modernes	3344			4,67		
Philosophie/Pédagogie/Psychologie	1207			4,67		
Musique	729				4,75	
Langues anciennes	682					4,87

$d_{\max} = 0,60$

L'amplitude maximale de l'effet est moyenne lorsqu'on compare les notes de maturité de langue première (voir tableau IV.61). A noter que les bachelières et bacheliers qui ont opté pour l'option spécifique «Musique» s'en sortent mieux qu'aux tests, alors que la sous-population «Biologie et chimie» perd du terrain.

Tableau IV.62: Comparaison des notes de maturité en mathématiques selon l'option spécifique choisie

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Philosophie/Pédagogie/Psychologie	1207	4,14				
Arts visuels	1152	4,19				
Langues modernes	3344		4,26			
Economie et droit	3138		4,28			
Musique	729		4,30	4,31		
Biologie et chimie	1713			4,34		
Langues anciennes	682				4,57	
Physique et applications des mathématiques	1421					4,88

$d_{\max} = 0,93$

En mathématiques, l'amplitude maximale de l'effet est deux fois moins grand lorsque l'on compare les notes de maturité plutôt que les résultats des tests, mais il reste très élevé (voir tableau IV.62). La hiérarchie entre les sous-populations ne subit pour ainsi dire aucun changement. En mathématiques, les examinatrices et examinateurs arrivent apparemment plus facilement à s'en tenir des standards de qualité objectifs lorsqu'ils délivrent des notes, même si ces standards ne sont qu'implicites. Les bachelières et bacheliers de la sous-population «Langues anciennes» se distinguent une nouvelle fois en damant le pion aux élèves qui ont choisi l'option spécifique «Biologie et chimie».

Tableau IV.63: Comparaison des notations continues comptant pour la maturité en biologie selon l'option spécifique choisie

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Musique	692	4,61			
Langues modernes	3007	4,61			
Economie et droit	2635	4,65	4,65		
Philosophie/Pédagogie/Psychologie	985		4,66		
Physique et applications des mathématiques	1033		4,69		
Biologie et chimie	1082			4,80	
Arts visuels	1056			4,81	
Langues anciennes	618				4,91

$d_{\max} = 0,54$

En biologie, l'amplitude maximale de l'effet s'abaisse à un niveau moyen lorsque l'on compare les notes obtenues durant la dernière année (voir tableau IV.63). A quelques petites différences près, la hiérarchie entre les sous-populations est étonnamment similaire à celle qui découle du test de biologie. Les plus grandes différences concernent les bachelières et bacheliers des options spécifiques «Arts visuels» et «Biologie et chimie». Les premiers gagnent du terrain. Les seconds en perdent, se faisant dépasser par la sous-population «Langues anciennes», ce qui est plutôt singulier.

Tableau IV.64: Comparaison des notes de maturité en sciences expérimentales selon l'option spécifique choisie

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Musique	729	4,48				
Philosophie/Pédagogie/Psychologie	1207	4,49				
Langues modernes	3344	4,52				
Economie et droit	3138		4,58			
Arts visuels	1152			4,65		
Biologie et chimie	1691			4,65		
Physique et applications des mathématiques	1423				4,72	
Langues anciennes	682					4,83

d max = 0,64

La note finale de sciences expérimentales ne peut être comparée avec les résultats du test de biologie que sous réserves, mais la hiérarchie est quand même intéressante (voir tableau IV.64). L'amplitude maximale de l'effet est nettement plus basse que lors du test de biologie. La meilleure sous-population et de loin est celle qui regroupe les élèves de l'option spécifique «Langues anciennes». Étonnamment, celle-ci s'avère plus performante que les spécialistes en sciences expérimentales sur leur propre terrain. La sous-population «Langues anciennes» est aussi celle qui obtient la meilleure notation continue en biologie. L'amplitude maximale de l'effet est un plus basse que pour la note finale en sciences expérimentales.

Le tableau IV.65 montre les pourcentages de résultats insuffisants enregistrés aux examens de maturité à titre de complément. En mathématiques, les bacheliers et bacheliers qui ont choisi l'une des options spécifiques créées avec le RRM 95 sont souvent à la peine. La moitié au moins des élèves des sous-populations «Philosophie/Pédagogie/Psychologie», «Arts visuels» et «Musique» ont obtenu des notes insuffisantes à l'écrit. Souvent meilleures, les notes attribuées en cours d'année par les enseignants ont manifestement un effet compensatoire par rapport au nombre plutôt élevé de notes insuffisantes aux examens écrits.

Tableau IV.65: Comparaison de la proportion de notes insuffisantes aux examens de maturité selon l'option spécifique choisie

Options spécifiques	Examen écrit de langue première	Note finale de maturité en langue première	Examen écrit de mathématiques	Note finale de maturité en mathématiques	Biologie: notation continue	Note finale de maturité en sciences expérimentales
Langues anciennes	9,7%	1,3%	32,3%	19,9%	6,0%	2,7%
Langues modernes	19,2%	4,3%	46,7%	29,6%	8,7%	8,4%
Physique et applications des mathématiques	14,9%	4,8%	12,4%	6,3%	6,1%	2,3%
Biologie et chimie	22,7%	6,2%	41,8%	17,2%	3,2%	3,5%
Economie et droit	22,0%	5,8%	40,7%	25,5%	4,8%	5,2%
Philosophie/Pédagogie/Psychologie	20,6%	3,7%	55,8%	32,3%	3,4%	8,9%
Arts visuels	24,2%	5,3%	48,8%	32,3%	2,5%	3,4%
Musique	13,0%	1,3%	48,6%	25,6%	5,4%	5,9%

6.3.3 Résultats du test des compétences transversales

Les quatre prochains tableaux (IV.66 à IV.69) comparent les compétences transversales qui ont été évaluées dans le cadre du TCT entre les sous-populations.

La hiérarchie entre les sous-populations est identique dans tous les domaines. Les différences entre les meilleurs («Physique et applications mathématiques») et les plus mauvais («Philosophie/Pédagogie/Psychologie») sont considérables (amplitude de l'effet très important). Les bachelières et bacheliers de l'option spécifique «Arts visuels» font également partie du groupe significativement le plus faible.

Tableau IV.66: Comparaison des résultats du test des compétences transversales selon l'option spécifique choisie (TCT, quantités et formes)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Philosophie/Pédagogie/Psychologie	834	469				
Arts visuels	901	473				
Musique	665		485			
Langues modernes	2426		486			
Economie et droit	2359			499		
Langues anciennes	391				514	
Biologie et chimie	939				517	
Physique et applications des mathématiques	761					560

d max = 1,25

Tableau IV.67: Comparaison des résultats du test des compétences transversales selon l'option spécifique choisie (TCT, diagrammes et tableaux)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5, 6)					
		1	2	3	4	5	6
Philosophie/Pédagogie/Psychologie	834	468					
Arts visuels	901	475	475				
Musique	665		481	481			
Langues modernes	2426			487			
Economie et droit	2359				499		
Langues anciennes	391				506		
Biologie et chimie	939					519	
Physique et applications des mathématiques	761						557

d max = 1,22

Tableau IV.68: Comparaison des résultats du test des compétences transversales selon l'option spécifique choisie (TCT, compréhension des textes)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Philosophie/Pédagogie/Psychologie	834	476				
Arts visuels	901	478				
Musique	665	483	483			
Langues modernes	2426		488			
Economie et droit	2359			496		
Langues anciennes	391				514	
Biologie et chimie	939				519	
Physique et applications des mathématiques	761					553

d max = 1,02

Tableau IV.69: Comparaison des résultats du test des compétences transversales selon l'option spécifique choisie (TCT, notions élémentaires de sciences expérimentales)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5, 6)					
		1	2	3	4	5	6
Philosophie/Pédagogie/Psychologie	834	474					
Arts visuels	901	477					
Musique	665		485				
Langues modernes	2426		488				
Economie et droit	2359			496			
Langues anciennes	391				511		
Biologie et chimie	939					521	
Physique et applications des mathématiques	761						549

d max = 1,01

6.3.4 Résultats choisis issus du questionnaire complémentaire

Les tableaux IV.70 à IV.74 montrent la façon dont les élèves des différentes sous-populations ont évalué leurs propres compétences transversales dans le questionnaire complémentaire. Seules les compétences pour lesquelles la différence entre les moyennes extrêmes est au moins de faible amplitude sont listées ($d > 0.2$).

Tableau IV.70: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (planification du temps de travail: 1 = mauvaise, 5 = très bonne)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Physique et applications des mathématiques	1418	2,32				
Biologie et chimie	1733		2,58			
Langues anciennes	656		2,63	2,63		
Langues modernes	3304			2,66	2,66	
Economie et droit	3145			2,69	2,69	
Musique	706			2,70	2,70	2,70
Arts visuels	1135				2,73	2,73
Philosophie/Pédagogie/Psychologie	1296					2,78

$d_{\max} = 0,50$

Les bacheliers et bacheliers ayant choisi l'option spécifique «Physique et applications des mathématiques» sont les plus critiques sur leur capacité de planifier leur temps de travail. La sous-population «Philosophie/Pédagogie/Psychologie» s'accorde les meilleures notes en la matière (voir tableau IV.70). L'écart entre les deux sous-populations est moyen. En chiffres absolus, il représente tout de même un demi-point sur l'échelle (qui va de 1 à 4). Les «spécialistes» en physique et en mathématiques sont également les plus réservés sur leur capacité de préparer efficacement les examens (voir tableau IV.71).

Tableau IV.71: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (efficacité de la préparation aux examens: 1 = inefficace, 5 = très efficace)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Physique et applications des mathématiques	1418	3,55				
Musique	706	3,59	3,59			
Economie et droit	3145		3,61	3,61		
Arts visuels	1135		3,62	3,62		
Biologie et chimie	1733			3,64		
Philosophie/Pédagogie/Psychologie	1296				3,72	
Langues modernes	3304				3,74	
Langues anciennes	656					3,86

$d_{\max} = 0,46$

Les différences présentées dans les tableaux suivants sont toutes significatives, mais de faible amplitude, et minimales en chiffres absolus.

Tableau IV.72: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (gestion du temps: 1 = très mauvaise, 5 = très bonne)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
Langues modernes	3304	3,67		
Musique	706		3,77	
Biologie et chimie	1733		3,77	
Philosophie/Pédagogie/Psychologie	1296		3,78	
Arts visuels	1135		3,78	
Economie et droit	3145		3,80	3,80
Langues anciennes	656		3,82	3,82
Physique et applications des mathématiques	1418			3,83

d max = 0,27

Tableau IV.73: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (capacité d'analyse et de synthèse: 1 = très faible, 5 = très développée)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4, 5)				
		1	2	3	4	5
Langues modernes	3299	3,20				
Physique et applications des mathématiques	1418		3,27			
Economie et droit	3145		3,29	3,29		
Biologie et chimie	1733		3,30	3,30		
Arts visuels	1135		3,30	3,30		
Musique	706			3,35	3,35	
Langues anciennes	655				3,40	3,40
Philosophie/Pédagogie/Psychologie	1296					3,44

d max = 0,31

Tableau IV.74: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (capacité d'adopter d'autres points de vue en cas de conflit: 1 = très bonne, 4 = très mauvaise)

Options spécifiques	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
Physique et applications des mathématiques	1409	2,92			
Economie et droit	3139	2,92			
Biologie et chimie	1733		2,98		
Langues modernes	3295		3,01	3,01	
Philosophie/Pédagogie/Psychologie	1296			3,05	3,05
Musique	705			3,05	3,05
Arts visuels	1134				3,06
Langues anciennes	656				3,07

d max = 0,25

6.4 Comparaisons selon le sexe

Comme c'est le cas dans la plupart des études, les données principales réunies dans le cadre d'EVAMAR II (résultats de tests, notes de maturité, résultats de l'auto-évaluation des élèves) ont été analysées aussi sous l'angle de la différence entre les sexes.

6.4.1 Résultats de langue première

Le tableau IV.75 compare les performances des deux sexes en langue première. Dans la plupart des cas, les bacheliers enregistrent des performances significativement meilleures, mais les différences sont insignifiantes dans les résultats des tests. En revanche, elles ont obtenues de meilleures notes à l'examen écrit de maturité (petite Amplitude de l'effet). L'écart avec les bacheliers est encore un peu plus élevé sur les notes obtenues en classe durant la dernière année de cours.

Tableau IV.75: Ecart en langue première selon le sexe

Types de résultats et d'examen	Sexe	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_LANGUE	Masculin	5539	498	90,21	-4	0,009	0,05
	Féminin	7822	502	82,96			
C-E_GEN	Masculin	5539	499	85,08	-2	0,140	0,03
	Féminin	7822	501	75,63			
C-E_DETAIL	Masculin	5539	495	89,39	-9	0,000	0,11
	Féminin	7822	504	80,86			
VOC	Masculin	5539	503	76,18	4	0,003	0,05
	Féminin	7822	499	70,22			
GRAM	Masculin	5539	497	78,07	-6	0,000	0,08
	Féminin	7822	503	71,06			
Langue première: notation continue	Masculin	5573	4,49	.520	-0,20	0,000	0,38
	Féminin	7824	4,69	.498			
Langue première: oral	Masculin	5563	4,60	.760	-0,10	0,000	0,13
	Féminin	7820	4,70	.750			
Langue première: écrit	Masculin	5571	4,37	.750	-0,07	0,000	0,10
	Féminin	7822	4,44	.730			
Langue première: note finale	Masculin	5569	4,55	.547	-0,14	0,000	0,25
	Féminin	7820	4,69	.532			

6.4.2 Résultats de mathématiques

Le tableau IV.76 compare les performances des deux sexes en mathématiques. Les bacheliers s'en sortent mieux que les bachelères dans tous les domaines examinés. Les écarts sont d'un niveau moyen pour les tests. Pour les notes de maturité, les différences sont faibles (examen écrit) à insignifiantes (examen oral et notation continue).

Tableau IV.76: Ecart en mathématiques selon le sexe

Types de résultats et d'examen	Sexe	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_MATH	Masculin	5666	523	96,26	49	.000	0,56
	Féminin	7858	474	81,76			
ANALYSE	Masculin	5666	523	94,84	47	.000	0,54
	féminin	7858	476	77,11			
GEOMETRIE	masculin	5666	525	91,94	49	.000	0,58
	féminin	7858	476	76,35			
STOCHASTIQUE	masculin	5666	519	88,20	38	.000	0,45
	féminin	7858	481	81,77			
DIVERS_M	masculin	5666	515	86,13	35	.000	0,42
	féminin	7858	480	78,20			
Mathématiques: notation continue	masculin	5573	4,42	.787	.08	.000	0,11
	féminin	7808	4,34	.795			
Mathématiques: oral	masculin	5091	4,43	1,385	.10	.000	0,08
	féminin	7394	4,33	1,007			
Mathématiques: écrit	masculin	5570	4,19	1,052	.27	.000	0,25
	féminin	7811	3,92	1,032			
Mathématiques: note finale	masculin	5569	4,42	.830	.13	.000	0,16
	Féminin	7820	4,29	.819			

6.4.3 Résultats de biologie

Le tableau IV.77 compare les performances des deux sexes en biologie et, pour ce qui est des notes de maturité, dans d'autres domaines des sciences expérimentales. Les bachelières enregistrent de meilleurs résultats dans tous les domaines du test de biologie, mais les écarts sont faibles. La note finale de maturité en sciences expérimentales n'est pas significativement différente entre les sexes. Pour ce qui est des notes obtenues en classe durant la dernière année, les bachelières sont meilleures en biologie et les bacheliers en physique. En chimie, les différences entre les deux sexes sont insignifiantes. Les bachelières compensent manifestement leurs mauvaises notes en physique par de meilleures notes en biologie.

Tableau IV.77: Ecart en biologie et en sciences naturelles selon le sexe

Types de résultats et d'examen	Sexe	N (pondéré)	Moyenne	Ecart type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_BIOL	masculin	5573	507	87,56	17	0,000	0,19
	féminin	7868	490	81,42			
B_EVO	masculin	5573	507	81,54	16	0,000	0,19
	féminin	7868	491	76,41			
B_STRU	masculin	5573	507	84,08	16	0,000	0,19
	féminin	7868	491	77,60			
B_GEN	masculin	5573	506	83,49	14	0,000	0,17
	féminin	7868	492	78,32			
B_INFO	masculin	5573	506	84,01	15	0,000	0,19
	féminin	7868	491	78,74			
B_ECO	masculin	5573	506	84,50	16	0,000	0,20
	féminin	7868	490	78,90			
B_MET	masculin	5573	506	86,06	15	0,000	0,18
	féminin	7868	491	79,46			
Biologie: notation continue	masculin	4460	4,63	0,540	-.09	0,000	0,18
	féminin	6652	4,72	0,539			
Chimie: notation continue	masculin	4511	4,60	0,633	.05	0,000	0,08
	féminin	6715	4,55	0,603			
Physique: notation continue	masculin	4413	4,59	0,684	.16	0,000	0,23
	féminin	6740	4,43	0,697			
Sciences expérimentales: note finale	masculin	5565	4,60	0,575	.01	0,147	0,03
	féminin	7805	4,59	0,560			
Physique (OS): note finale	masculin	1062	4,41	0,887	-.15	0,014	0,18
	féminin	263	4,56	0,780			

6.4.4 Résultats du test des compétences transversales

Les différences entre les sexes ne sont nulle part aussi importantes que dans le test des compétences transversales (TCT; voir tableau IV.78). L'amplitude de l'effet y est invariablement de l'ordre d'un effet moyen. Cet écart (en faveur des bacheliers) s'explique sans doute principalement du fait que les filles sont proportionnellement plus nombreuses à suivre une formation gymnasiale que les garçons d'une même tranche d'âge en Suisse, pour un niveau moyen d'aptitudes sensiblement égal. Le nombre plus important de bachelières

au sommet de la pyramide des aptitudes (également distribuées entre les sexes) se répercute donc négativement sur leur niveau moyen d'aptitudes.

Tableau IV.78: Ecart dans les résultats du TCT selon le sexe

Types de résultats	Sexe	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
T_QUANT	masculin	3833	525	82,60	48	.000	0,59
	féminin	5448	477	78,84			
T_DIAGR	masculin	3833	524	79,91	47	.000	0,60
	féminin	5448	477	78,08			
T_TEXT	masculin	3833	521	82,47	41	.000	0,51
	féminin	5448	480	78,19			
T_SCIEN	masculin	3833	521	82,04	41	.000	0,51
	féminin	5448	480	76,81			

6.4.5 Résultats choisis issus du questionnaire complémentaire

Le tableau IV.79 montre la façon dont les élèves des deux sexes ont évalué leurs propres compétences transversales dans le questionnaire complémentaire. Seules les compétences pour lesquelles la différence entre les moyennes extrêmes est au moins de faible amplitude sont listées ($d > 0.2$). Les bachelières indiquent être plus aptes à planifier leur temps de travail et à préparer efficacement leurs examens. Elles disent aussi être plus promptes à adopter le point de vue d'autres membres d'un groupe dans des situations conflictuelles. Les bacheliers pensent quant à eux qu'ils gèrent plus efficacement leur temps.

Tableau IV.79: Domaines du TCT où les écarts sont significatifs entre les sexes (amplitude de l'effet $> 0,2$)

Domaines de compétence	Sexe	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
Planification du temps de travail	masculin	5540	2,41	.944	-.39	.000	0,42
	féminin	7856	2,80	.939			
Efficacité de la préparation des examens	masculin	5540	3,49	.697	-.29	.000	0,45
	féminin	7856	3,78	.614			
Gestion du temps	masculin	5540	3,88	.648	.20	.000	0,32
	féminin	7856	3,68	.631			
Capacité d'adopter le point de vue d'autrui	masculin	5523	2,90	.612	-.15	.000	0,27
	féminin	7847	3,05	.544			

6.5 Comparaisons selon le taux de maturités

Les taux de maturités varient de façon considérable en Suisse. Cela ne traduit toutefois pas nécessairement une inégalité des chances dans l'accès à la maturité supérieure. Mais il serait évidemment intéressant de savoir si les exigences fixées à l'obtention de la maturité sont les mêmes partout. Divers effets s'enchevêtrent en l'occurrence. Pour pouvoir juger si l'égalité des chances est respectée, il faudrait tenir compte déjà de la sélection qui se fait à l'entrée des gymnases. Certains déficits de compétence peuvent tout à fait être «rattrapés» durant le gymnase. Mais l'inégalité des chances qui existait au départ subsiste néanmoins. Lors de l'année scolaire 2006/2007, 29,5% des jeunes adultes ont décroché la maturité au Tessin contre 11,2% seulement à Obwald (OFS, 2008b). EVAMAR II permet de comparer entre elles certaines compétences des bachelières et bacheliers issus de cantons ayant des taux de maturités différents. L'échantillon n'étant pas représentatif dans tous les cantons, les

cantons ont été regroupés grossièrement en trois blocs de taille aussi égale que possible d'après leurs taux de maturités. Le tiers inférieur comprend les cantons où le taux de maturités est inférieur à 17,5% et le tiers supérieur ceux où il dépasse 19,0%. Les résultats des comparaisons sont exposés dans les tableaux IV.80 à IV.83.

Tableau IV.80: Ecart en langue première selon le taux de maturité dans les groupes de cantons

Types de résultats et d'examens	Taux de maturité	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_LANGUE	tiers inférieur	5483	508	83,69	12	.000	0,14
	tiers supérieur	4490	496	90,77			
C-E_GEN	tiers inférieur	5483	506	77,42	9	.000	0,12
	tiers supérieur	4490	497	83,44			
C-E_DETAIL	tiers inférieur	5483	510	81,30	17	.000	0,19
	tiers supérieur	4490	493	89,81			
VOC	tiers inférieur	5483	503	72,72	-2	.132	0,03
	tiers supérieur	4490	505	75,93			
GRAM	tiers inférieur	5483	505	77,00	7	.000	0,10
	tiers supérieur	4490	498	70,64			
Langue première: notation continue	tiers inférieur	5849	4,67	.469	.16	.000	0,31
	tiers supérieur	4497	4,51	.559			
Langue première: oral	tiers inférieur	5849	4,72	.721	.15	.000	0,20
	tiers supérieur	4487	4,57	.789			
Langue première: écrit	tiers inférieur	5849	4,50	.685	.20	.000	0,28
	tiers supérieur	4493	4,30	.784			
S_MatPr	tiers inférieur	5856	4,61	.614	.16	.000	0,25
	tiers supérieur	4487	4,45	.666			
Langue première: note finale	tiers inférieur	5849	4,70	.499	.16	.000	0,30
	tiers supérieur	4493	4,54	.572			

Tableau IV.81: Ecart en mathématiques selon le taux de maturité dans les groupes de cantons

Types de résultats et d'examens	Taux de maturité	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_MATH	tiers inférieur	5673	504	89,73	24	.000	0,26
	tiers supérieur	4404	480	92,71			
ANALYSE	tiers inférieur	5673	501	88,31	16	.000	0,18
	tiers supérieur	4404	485	88,06			
GEOMETRIE	tiers inférieur	5673	502	85,95	13	.000	0,15
	tiers supérieur	4404	489	87,76			
STOCHASTIQUE	tiers inférieur	5673	502	87,19	17	.000	0,19
	tiers supérieur	4404	485	84,06			
DIVERS_M	tiers inférieur	5673	507	83,72	30	.000	0,35
	tiers supérieur	4404	477	82,92			
Mathématiques: notation continue	tiers inférieur	5843	4,42	.765	.15	.000	0,20
	tiers supérieur	4497	4,27	.800			
Mathématiques: oral	tiers inférieur	5066	4,39	1,016	-.01	.706	0,01
	tiers supérieur	4371	4,40	1,410			
Mathématiques: écrit	tiers inférieur	5840	4,07	1,048	.12	.000	0,11
	tiers supérieur	4492	3,95	1,045			
Mathématiques: note finale	tiers inférieur	5849	4,37	.823	.09	.000	0,11
	tiers supérieur	4493	4,28	.810			

Tableau IV.82: Ecart en sciences expérimentales selon le taux de maturité dans les groupes de cantons

Types de résultats et d'examens	Taux de maturité	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_BIOL	tiers inférieur	5564	517	76,85	50	.000	0,61
	tiers supérieur	4462	467	86,02			
B_EVOL	tiers inférieur	5564	515	70,63	44	.000	0,58
	tiers supérieur	4462	471	81,35			
B_STRU	tiers inférieur	5564	516	72,77	46	.000	0,59
	tiers supérieur	4462	470	83,12			
B_GEN	tiers inférieur	5564	515	73,03	43	.000	0,55
	tiers supérieur	4462	472	83,32			
B_INFO	tiers inférieur	5564	516	72,78	47	.000	0,60
	tiers supérieur	4462	469	84,35			
B_ECO	tiers inférieur	5564	515	73,45	46	.000	0,57
	tiers supérieur	4462	469	85,58			
B_MET	tiers inférieur	5564	516	73,49	47	.000	0,59
	tiers supérieur	4462	469	86,04			
Biologie: notation continue	tiers inférieur	5841	4,67	.528	-.08	.000	0,15
	tiers supérieur	2377	4,75	.556			
Chimie: notation continue	tiers inférieur	5839	4,55	.597	-.01	.632	0,01
	tiers supérieur	2492	4,56	.654			
Physique: notation continue	tiers inférieur	5841	4,46	.689	-.03	.060	0,04
	tiers supérieur	2479	4,49	.718			
Sciences expérimentales: note finale	tiers inférieur	5825	4,62	.572	.08	.000	0,13
	tiers supérieur	4493	4,54	.561			

Tableau IV.83: Ecart dans le TCT selon le taux de maturité dans les groupes de cantons

Types de résultats et d'examens	Taux de maturité	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
T_QUANT	tiers inférieur	5162	503	82,93	15	.000	0,19
	tiers supérieur	832	488	78,69			
T_DIAGR	tiers inférieur	5162	503	81,78	17	.000	0,22
	tiers supérieur	832	486	75,28			
T_TEXT	tiers inférieur	5162	504	81,65	21	.000	0,26
	tiers supérieur	832	483	76,27			
T_SCIEN	tiers inférieur	5162	504	80,39	19	.000	0,25
	tiers supérieur	832	485	74,35			

Les bacheliers et bacheliers provenant du groupe où les taux de maturités sont les plus bas obtiennent des résultats légèrement meilleurs à tous les tests d'EVAMAR II. Les écarts sont les plus importants en biologie (amplitude de l'effet au moins moyenne). Le groupe où les taux de maturités sont les plus bas est également plus performant aux examens de maturité, mais de manière moins nette qu'aux tests. En langue première, la situation est toutefois inversée: les différences sont plus marquées sur les notes que sur les résultats des tests. Dans les grandes lignes, ces constats demeuraient vrais si les groupes de cantons avaient été découpés de la façon proposée en 2008 par Oelkers (il distinguait les cantons où le taux de maturités est inférieur à 15%, ceux où il dépasse 20% et les cantons intermédiaires). Globalement, les élèves provenant des régions où les taux de maturités sont plus élevés ont réalisé de moins bons résultats aux tests d'EVAMAR II. La question de l'égalité des chances se pose donc.

6.6 Comparaisons selon la durée de la voie gymnasiale

Une partie des gymnases suisses considèrent que la formation gymnasiale au sens large débute immédiatement après l'école primaire, autrement dit avec le niveau secondaire I. Une comparaison des résultats entre les établissements d'après la durée de la voie gymnasiale s'imposait donc (tableaux IV.84 à IV.87).

Tableau IV.84: Ecart en langue première selon la durée de la voie gymnasiale

Types de résultats et d'examens	Durée de la voie gymnasiale	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_LANGUE	Courte	11365	500	86,62	-5	.008	0,07
	Longue	1996	505	82,68			
C-E_GEN	Courte	11365	499	79,99	-9	.000	0,12
	Longue	1996	508	77,50			
C-E_DETAIL	Courte	11365	499	85,18	-12	.000	0,14
	Longue	1996	511	80,75			
VOC	Courte	11365	501	72,78	0	.998	0,00
	Longue	1996	501	72,78			
GRAM	Courte	11365	498	73,11	-16	.000	0,22
	Longue	1996	514	78,03			
Langue première: notation continue	Courte	11264	4,59	.523	-.10	.000	0,19
	Longue	2133	4,69	.475			
Langue première: oral	Courte	11254	4,66	.757	-.02	.212	0,03
	Longue	2129	4,68	.746			
Langue première: écrit	Courte	11260	4,39	.745	-.14	.000	0,19
	Longue	2133	4,53	.695			
Langue première: note finale	Courte	11260	4,62	.546	-.07	.000	0,14
	Longue	2129	4,69	.519			

Tableau IV.85: Ecart en mathématiques selon la durée de la voie gymnasiale

Types de résultats et d'examens	Durée de la voie gymnasiale	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_MATH	Courte	11432	493	92,61	-10	.000	0,11
	Longue	2092	503	84,59			
ANALYSE	Courte	11432	494	88,33	-9	.000	0,11
	Longue	2092	503	86,46			
GEOMETRIE	Courte	11432	496	87,62	-3	.109	0,04
	Longue	2092	499	81,01			
STOCHASTIQUE	Courte	11432	495	87,13	-10	.000	0,11
	Longue	2092	505	83,37			
DIVERS_M	Courte	11432	493	84,31	-13	.000	0,16
	Longue	2092	506	77,47			
Mathématiques: notation continue	Courte	11258	4,37	.791	-.03	.114	0,04
	Longue	2123	4,40	.802			
Mathématiques: oral	Courte	10356	4,37	1,208	-.01	.527	0,02
	Longue	2129	4,38	1,014			
Mathématiques: écrit	Courte	11249	4,01	1,051	-.13	.000	0,12
	Longue	2132	4,14	1,025			
Mathématiques: note finale	Courte	11260	4,33	.825	-.05	.012	0,06
	Longue	2129	4,38	.831			

Tableau IV.86: Ecart en sciences expérimentales selon la durée de la voie gymnasiale

Types de résultats et d'examens	Durée de la voie gymnasiale	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
TOT_BIOL	Courte	11433	493	84,51	-27	.000	0,32
	Longue	2008	520	80,18			
B_EVO	Courte	11433	494	78,86	-24	.000	0,31
	Longue	2008	518	76,34			
B_STRU	Courte	11433	494	80,80	-24	.000	0,30
	Longue	2008	518	77,19			
B_GEN	Courte	11433	494	80,88	-24	.000	0,30
	Longue	2008	518	77,21			
B_INFO	Courte	11433	493	81,64	-25	.000	0,31
	Longue	2008	518	76,15			
B_ECO	Courte	11433	493	81,80	-25	.000	0,31
	Longue	2008	518	77,55			
B_MET	Courte	11433	493	82,96	-25	.000	0,31
	Longue	2008	518	76,90			
Biologie: notation continue	Courte	8979	4,70	.539	.09	.000	0,17
	Longue	2133	4,61	.546			

Types de résultats et d'examens	Durée de la voie gymnasiale	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
Chimie: notation continue	Courte	9093	4,56	.612	-.05	.002	0,08
	Longue	2133	4,61	.628			
Physique: notation continue	Courte	9020	4,49	.701	-.05	.003	0,07
	Longue	2133	4,54	.673			
Sciences expérimentales: note finale	Courte	11236	4,59	.565	-.03	.077	0,04
	Longue	2133	4,62	.574			

Tableau IV.87: Ecart dans le TCT selon la durée de la voie gymnasiale

Types de résultats et d'examens	Durée de la voie gymnasiale	N (pondéré)	Moyenne	Ecart-type	Différence moyenne	Degré de signification	Amplitude de l'effet
T_QUANT	Courte	7196	495	82,14	-8	.000	0,09
	Longue	2085	503	88,99			
T_DIAGR	Courte	7196	495	81,38	-7	.000	0,09
	Longue	2085	502	84,73			
T_TEXT	Courte	7196	496	81,07	-6	.001	0,08
	Longue	2085	502	86,87			
T_SCIEN	Courte	7196	495	79,91	-7	.002	0,08
	Longue	2085	502	86,53			

Les élèves engagés dans une voie gymnasiale longue obtiennent des résultats significativement meilleurs aux examens de maturité et dans presque tous les domaines testés dans EVAMAR II. L'amplitude de l'effet est toutefois très faible. Elle ne dépasse 0,2 qu'au test de biologie (mais pas pour les notes) et dans le volet «Compétence grammaticale et orthographe» du test de langue première. Il est vraisemblable que les élèves reçoivent des bases plus solides en grammaire et en orthographe lorsque la voie gymnasiale est longue. Il est possible aussi que la formation en biologie soit conçue de façon plus systématiquement et débouche sur des connaissances plus étoffées quand la voie gymnasiale commence dès le niveau secondaire I. Dans les établissements où la voie gymnasiale est courte, le corps enseignant doit souvent commencer par reposer les bases élémentaires de biologie car les élèves n'auront pas tous acquis au niveau secondaire I des connaissances équivalentes. Les déperditions liées au passage du niveau secondaire I au niveau secondaire II sont probablement plus faibles dans les disciplines centrales comme les mathématiques et la langue première. Il n'est pas nécessaire de «tout reprendre à zéro» au gymnase car les compétences acquises par les élèves au niveau secondaire I sont plus uniformes dans ces matières. Les notes délivrées en classe durant la dernière année de cours (voir tableau IV.86) ont une particularité intéressante. C'est le seul domaine de maturité où les élèves qui fréquentent des «gymnases courts» s'en sortent mieux. A priori contradictoire, ce résultat montre en fait de manière particulièrement claire que la pratique d'évaluation dans les gymnases a tendance à s'adapter au niveau de la classe quand il n'existe pas d'exigences standardisées.

6.7 Autres comparaisons

6.7.1 Assiduité et résultats des tests

Les bacheliers et bacheliers devaient dire dans le questionnaire complémentaire s'ils avaient participé aux tests avec sérieux. Leur réponse à l'affirmation «Je me suis appliqué» a été enregistrée sur une échelle à quatre niveaux («pas du tout», «plutôt pas», «plutôt», «tout à fait»). Les résultats montrent que ceux qui se sont appliqués ont aussi mieux réussi les tests. Le tableau IV.88 l'illustre pour les mathématiques à titre d'exemple.

Tableau IV.88: Lien entre l'application des élèves lors des tests et leurs résultats de mathématiques

Degré d'application	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
pas du tout	553	477		
plutôt pas	2791	478		
plutôt	7686		496	
tout à fait	1905			522

d max = 0,48

Le fait que les résultats dépendent de l'assiduité des élèves semble mettre en cause la validité des tests d'EVAMAR au premier abord. Mais l'assiduité des élèves a aussi été testée par rapport à leurs notes de maturité. La hiérarchie des notes reste la même que pour les tests d'EVAMAR dans toutes les disciplines. Les tableaux IV.89 à IV.91 l'illustrent de nouveau pour les mathématiques à titre d'exemple. Les bacheliers et bacheliers peu assidus lors des tests obtiennent aussi de moins bonnes notes à l'école. On peut en conclure que le manque d'application de certains élèves n'a pas biaisé les résultats des tests, du moins pas de manière déterminante.

Tableau IV.89: Lien entre l'application des élèves lors des tests et leur notations continues en mathématiques

Degré d'application	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
pas du tout	570	4,14		
plutôt pas	2705	4,15		
plutôt	7546		4,42	
tout à fait	1867			4,60

d max = 0,60

Tableau IV.90: Lien entre l'application des élèves lors des tests et leurs notes de maturité à l'examen écrit en mathématiques

Degré d'application	N (pondéré)	Groupes significatifs (1, 2, 3, 4)			
		1	2	3	4
pas du tout	568	3,73			
plutôt pas	2705		3,83		
plutôt	7557			4,09	
tout à fait	1866				4,28

d max = 0,53

Tableau IV.91: Lien entre l'application des élèves lors des tests et leurs notes finales en mathématiques

Degré d'application	N (pondéré)	Groupes significatifs (1, 2, 3)		
		1	2	3
pas du tout	568	4,12		
plutôt pas	2705	4,15		
plutôt	7558		4,38	
tout à fait	1866			4,58

d max = 0,57

6.7.2 Examens oraux et écrits

Dans certaines disciplines, l'obtention de la maturité passe par des examens écrits et oraux. Il est donc intéressant d'examiner si ces deux types de notes ont un lien et s'il existe des différences significatives entre elles. Nous avons calculé la corrélation entre ces deux types de notes pour les trois disciplines testées dans EVAMAR II puis utilisé un test t pour échantillons appariés pour vérifier la signification des différences. Les résultats figurent dans le tableau IV.92.

Tableau IV.92: Comparaison des notes obtenues à l'examen oral et écrit

Notes	Moyenne	N (pondéré)	Différence de moyenne	Degré de signification	Corrélation
Mathématiques: oral	4,36	12474	.31	.000	.608
Mathématiques: écrit	4,05	12474			
Langue première: oral	4,66	13383	.25	.000	.436
Langue première: écrit	4,41	13383			
Biologie: oral	4,93	171	.45	.000	.489
Biologie: écrit	4,48	171			
Physique et maths (OS): oral	4,50	1259	.25	.000	.537
Physique et maths (OS): écrit	4,25	1259			
Biologie et chimie (OS): oral	4,57	1499	.41	.000	.504
Biologie et chimie (OS): écrit	4,16	1499			
Mathématiques (OC): oral	4,36	91	.02	.726	.808
Mathématiques (OC): écrit	4,34	91			
Biologie (OC): oral	5,24	129	.50	.000	.355
Biologie (OC): écrit	4,74	129			

Dans la plupart des cas, les corrélations sont modérées, hormis en mathématiques où r est élevé. Les examens oraux et écrit ne mesurent donc manifestement qu'en partie la même chose. Les notes de l'oral sont presque toujours meilleures que celles de l'écrit. Les notes de mathématiques en option complémentaire sont les seules où l'on ne constate aucune différence significative.

7 Corrélations

La batterie d'instruments utilisée dans le cadre d'EVAMAR II inclut diverses variables qui permettent de prédire le succès aux études dans une certaine mesure. Le test portant sur la langue première met l'accent sur des compétences linguistiques transversales qui sont essentielles quelles que soient les études poursuivies. Dans le test de mathématiques, les items portent sur des contenus qui, pour la plupart, s'avèreront importants dans une large palette de branches. Le test de biologie est axé davantage sur des connaissances spécialisées. Celles-ci seront pertinentes avant tout dans le cadre d'études en biologie ou en médecine. Elles revêtent aussi une certaine importance dans d'autres domaines que celui des sciences expérimentales. Le test des compétences transversales (TCT) sert à évaluer une grande variété de compétences, tout comme le test d'aptitude aux études de médecine, auquel il s'adosse. Il est donc intéressant d'examiner les relations entre les performances dans ces différents tests. Les corrélations entre elles seront examinées ci-après (point 7.1).

En Suisse, il suffit de détenir un certificat de maturité pour pouvoir être admis dans les universités, hormis dans les facultés de médecine. C'est pourquoi nous examineront ci-après les corrélations entre les résultats enregistrés aux tests de performance d'EVAMAR II et les notes de maturité (point 7.2).

7.1 Corrélations des tests entre eux

Le tableau IV.93 montre les corrélations qui existent entre les résultats globaux des différents tests. La valeur globale pour le TCT (TOT_TCT) a été également estimée pour la circonstance, en utilisant le modèle de Rasch.

Tableau IV.93: Corrélations entre les tests

Types de résultats et d'examens	TOT_MATH	TOT_BIOL	TOT_TCT
TOT_LANGUE	.148(**)	.224(**)	.298(**)
N (pondéré)	12964	13182	8740
TOT_MATH		.223(**)	.448(**)
N (pondéré)		12994	9082
TOT_BIOL			.362(**)
N (pondéré)			8860

** = significatif à un degré de 1%

Toutes les corrélations sont significatives, mais pas très élevées. La plus forte corrélation est celle entre le TCT et le test de mathématiques ($r = 0,448$). Ces résultats montrent que le projet EVAMAR II couvre différentes facettes de l'aptitude aux études supérieures, ce qui était son but.

Corrélations entre les tests et les notes de maturité

Les tableaux IV.94 à IV.96 montrent les corrélations qui existent entre les tests dans les différentes disciplines et les notes de maturité dans les branches qui s'y rattachent.

Tableau IV.94: Corrélations entre le test et les notes de maturité en langue première

Types de résultats et d'examens	Langue première: notation continue	Langue première: oral	Langue première: écrit	Langue première: note finale
TOT_LANGUE	.255(**)	.193(**)	.240(**)	.272(**)
N (pondéré)	12803	12789	12800	12796
Langue première: notation continue		.474(**)	.616(**)	.836(**)
N (pondéré)		13383	13394	13390
Langue première: oral			.436(**)	.732(**)
N (pondéré)			13383	13383
Langue première: écrit				.787(**)
N (pondéré)				13390

** = significatif à un degré de 1%

Tableau IV.95: Corrélations entre le test et les notes de maturité en mathématiques

Types de résultats et d'examens	Mathématiques: notation continue	Mathématiques: oral	Mathématiques: écrit	Mathématiques: note finale
TOT_MATH	.395(**)	.296(**)	.458(**)	.456(**)
N (pondéré)	12938	12092	12938	12947
Ma_ErfNote		.539(**)	.752(**)	.911(**)
N (pondéré)		12468	13365	13373
Mathématiques: oral			.608(**)	.698(**)
N (pondéré)			12474	12485
Mathématiques: écrit				.870(**)
N (pondéré)				13378

** = significatif à un degré de 1%

Tableau IV.96: Corrélations entre le test et les notes de maturité en biologie

Types de résultats et d'examens	Biologie: notation continue	Sciences expérimentales: note finale
TOT_BIOL	.250(**)	.257(**)
N (pondéré)	10610	12834
Biologie: notation continue		.714(**)
N (pondéré)		11088

** = significatif à un degré de 1%

Les corrélations entre les tests et les notes de maturité sont toutes significatives, mais le degré de corrélation n'est que moyen dans le meilleur des cas. Le fait que les notes ne soient pas standardisées explique en partie l'absence de corrélations plus élevées. En outre, le contenu des tests et celui des examens de maturité ne se recoupent qu'en partie. Le lien le plus fort existe en mathématiques ($r = 0,458$ pour l'examen écrit) et le plus faible en langue première. Cela correspond à ce que l'on pouvait attendre. Les aptitudes linguistiques testées

dans le cadre d'EVAMAR II ont en partie un caractère transversal. Elles ne sont donc pas comparables avec le contenu des cours de langue première (allemand ou français). Lors des examens de maturité, les questions portent sur des aspects beaucoup plus déterminants du point de vue de l'aptitude générale aux études (voir chapitre V). Il n'en reste pas moins que, globalement, c'est en langue première que les connaissances et les capacités testées à l'école s'éloignent le plus des compétences mesurées par les tests.

Le tableau IV.97 ci-après montre que le TCT est également corrélé de manière significative avec les notes de maturité, mais plus faiblement que les tests des trois disciplines auxquelles s'intéresse d'EVAMAR II (langue première, mathématiques, biologie).

Tableau IV.97: Corrélations entre le TCT et les notes de maturité

	Notation continue	Examen oral de maturité	Examen écrit de maturité	Note finale de maturité
TOT_TCT et langue première	.101(**)	.122(**)	.198(**)	.155(**)
N (pondéré)	8740	8735	8740	8735
TOT_TCT et mathématiques	.287(**)	.227(**)	.332(**)	.303(**)
N (pondéré)	8724	7947	8729	8735
TOT_TCT et biologie	.227(**)			
N (pondéré)	8501			
TOT_TCT et chimie	.277(**)			
N (pondéré)	8607			
TOT_TCT et chimie	.269(**)			
N (pondéré)	8580			
TOT_TCT et sciences expérimentales				.289(**)
N (pondéré)				8715

** = significatif à un degré de 1%

7.2 Corrélations des notes de maturité entre elles

Les notes sanctionnent potentiellement des performances très hétérogènes. Le tableau IV.98 montre les corrélations des notes de maturité entre elles dans les disciplines testées. Elles sont à nouveau significatives mais au mieux d'une intensité moyenne ($r = 0.582$ entre les sciences expérimentales et les mathématiques).

Tableau IV.98: Corrélations des notes de maturité entre elles

	Mathématiques: note finale	Biologie: notation continue	Sciences expérimentales: note finale
Langue première: note finale	.203(**)	.287(**)	.271(**)
N (pondéré)	13390	11108	13365
Mathématiques: note finale		.396(**)	.582(**)
N (pondéré)		11108	13365

** = significatif à un degré de 1%

8 Synthèse

8.1 Bases des tests et de l'enquête complémentaire

Un échantillon d'environ 3800 bacheliers et bacheliers alémaniques et romands a participé aux tests et à l'enquête complémentaire d'EVAMAR II. Les items ont été conçus à partir de supports de cours universitaires, de l'analyse des résultats des premiers examens intermédiaires dans les 16 branches les plus étudiées dans les universités suisses et des résultats d'une enquête menée auprès du corps professoral des institutions qui enseignent ces matières. En outre, il a été tenu compte des connaissances actuelles (théoriques et appliquées) dans la mesure des performances cognitives et dans la prédictibilité du succès des études supérieures.

Dans le test de langue première, l'accent a été mis sur les compétences linguistiques transversales, indispensables pour bien aborder des études supérieures dans n'importe quelle branche.

La maîtrise des contenus de la plupart des items du test de mathématiques est importante dans la perspective d'études dans un grand nombre de branches, et pas seulement en mathématiques.

Le test de biologie porte sur des connaissances spécialisées qui seront surtout utiles dans le cadre d'études supérieures en biologie et en médecine, mais aussi en partie hors du champ des sciences expérimentales.

Le test des compétences transversales (TCT) s'adosse au test d'aptitude aux études de médecine. Il mesure les compétences cognitives générales. Ce test n'a qu'une valeur complémentaire et il n'a été réalisé qu'en Suisse alémanique.

On dira d'une personne qu'elle dispose d'une aptitude générale aux études si elle est armée pour se lancer sans problèmes dans des études supérieures dans n'importe quelle branche. De ce point de vue, la batterie d'instruments utilisés dans EVAMAR II peut être qualifiée d'équilibrée, étant donné la diversité des variables prises en compte pour prédire le succès des études. Ces instruments ont été conçus en tenant compte de divers aspects de l'aptitude aux études décrits au début de ce chapitre: les capacités cognitives générales, les connaissances déjà acquises et l'aptitude à la lecture (voir section 1.2.1). Les compétences testées ne contribuent pas toutes à renforcer dans la même mesure l'aptitude générale aux études. Ainsi, une bonne maîtrise de la langue première est essentielle pour réussir dans pratiquement toutes les branches; des bases solides en mathématiques sont indispensables dans de nombreuses branches qui s'appuient sur le langage mathématique formel; des connaissances en biologie ne sont nécessaires que dans un nombre de branches limité; quant aux capacités mesurées à l'aide du TCT, elles sont significativement corrélées avec le succès d'études en médecine et dans d'autres branches qui exigent des compétences similaires. Les pré-analyses ont montré qu'il est avantageux pour de futurs étudiants d'acquérir dès le gymnase des connaissances dans la branche à laquelle ils se consacreront à l'université. A des degrés divers, cela était valable dans toutes les branches.

8.2 Résultats d'ensemble

Environ la moitié des items a été résolue correctement en moyenne. Le taux de bonnes réponses et le niveau des aptitudes personnelles estimées sont légèrement supérieurs en langue première qu'en mathématiques et en biologie. En biologie, les bacheliers et bacheliers qui avaient terminé leurs cours depuis au moins six mois ont été nettement moins performants au test, sans doute victimes de l'effet d'oubli qui s'instaure au fil du temps. Les connaissances emmagasinées au gymnase ne semblent acquises dans leur totalité que pendant une courte période. Il est donc souhaitable que la fin des cours au gymnase intervienne aussi tard que possible avant les examens de maturité et que les élèves entament rapidement leurs études supérieures dans la foulée. Les résultats du TCT sont légèrement moins bons que ceux obtenus au test réalisé chaque année pour mesurer l'aptitude des candidats aux études de médecine. Cet écart est plausible car des élèves qui estiment ne pas avoir les capacités pour entamer des études de médecine ont aussi participé aux tests d'EVAMAR. La faiblesse de l'écart est un indice fort que les bacheliers et bacheliers testés se sont donné beaucoup de peine pour répondre

aux questions qui leur étaient soumises. Leur engagement a été comparable à celui dont ils font preuve habituellement en classe lors d'épreuves sanctionnées par des notes. Dans le cas contraire, on ne s'expliquerait que difficilement que les résultats du TCT soient aussi proches de ceux réalisés dans le cadre d'un test qui conditionne aussi fortement l'avenir personnel que celui d'aptitude aux études de médecine.

Les différences de performance entre les élèves et entre les classes sont considérables. Les écarts sont plus importants qu'aux examens de maturité. Une part substantielle des bacheliers et des bachelères a vraisemblablement des compétences insuffisantes dans au moins un des domaines testés. L'analyse des résultats des examens de maturité – où les écoles délivrent elles-mêmes les notes – confirme ce point de vue, surtout en mathématiques. En 2007, 41,4% des élèves avaient obtenu une note insuffisante lors de l'examen écrit dans cette discipline. La note finale de mathématiques, qui tient compte aussi des notes obtenues en classe durant l'année et des résultats de l'examen oral, reste insuffisante dans 24,4% des cas. La question de savoir si les résultats insuffisants enregistrés aux examens de maturité trahissent des déficits de connaissances et de compétences importants sous l'angle de l'aptitude générale aux études reste posée. Dans des cas extrêmes, des bacheliers et des bachelères ayant des notes insuffisantes disposent peut-être quand même d'aptitudes suffisantes dans les domaines qui comptent; à l'inverse, certains élèves qui ont obtenu des notes de maturité suffisantes ne sont peut-être pas dotés de toutes les compétences nécessaires dans les domaines cruciaux du point de vue de leur aptitude générale aux études. En revanche, les résultats en partie médiocres obtenus aux tests trahissent clairement des déficits déterminants pour l'aptitude aux études dans au moins une discipline. Il n'est pas possible d'établir de façon déterministe à partir de quel niveau de performance aux tests des études supérieures couronnées de succès deviendraient illusoires, à moins de mener des enquêtes empiriques étendues et de longue haleine. Mais en tenant compte aussi des notes obtenues aux examens de maturité, on peut raisonnablement penser qu'une partie des bacheliers et des bachelères ne disposent pas de compétences suffisantes pour pouvoir se lancer dans des études supérieures dans n'importe quelle discipline – à moins qu'ils ne comblient leurs lacunes par ailleurs. Pourtant, le certificat de maturité est censé attester leur aptitude générale aux études. Les lacunes en mathématiques et en langue première ressenties par le corps professoral confirment ce point de vue (voir chapitre II, section 3.3). Cette situation paraît inévitable dans le système actuel, à cause de la possibilité de compenser les notes insuffisantes à certains examens par de meilleurs résultats obtenus ailleurs.

8.3 Résultats des tests du point de vue structurel

Les résultats des tests semblent dépendre de la durée des études au gymnase. Les établissements où la durée réglementaire minimale est de quatre ans obtiennent de meilleurs résultats que ceux où la première des quatre années peut être accomplie dans une école du niveau secondaire I dans le cadre d'un enseignement à caractère «prégymnasial». La durée du gymnase a un impact sur les résultats du test de mathématiques, à un faible degré sur les résultats du test de langue première et, en Suisse alémanique, sur les résultats du TCT.

Les gymnases de longue durée affichent de meilleurs résultats dans presque tous les domaines testés, mais la différence est petite. Le seul domaine où l'écart est plus marqué est la biologie, et ce bien que les notes délivrées au cours de la dernière année de cours dans cette discipline soient légèrement mais significativement plus élevées dans les établissements de courte durée. Cela dit, il s'agit probablement d'une anomalie (notation qui s'adapte au niveau de la classe).

Les résultats peuvent varier de manière considérable selon l'option spécifique (OS) choisie par les élèves. Certaines différences étaient attendues. Les élèves qui ont choisi l'OS «Physique et applications des mathématiques» ont par exemple brillé en mathématiques. La sous-population «Biologie et chimie» a obtenu d'excellents résultats en biologie. Certains écarts surprennent davantage. Les résultats des tests de langue première sont très bons chez les bacheliers et des bachelères qui ont choisi l'OS «Langues anciennes» mais seulement moyens chez ceux qui ont opté pour les langues modernes. Il n'y a pas d'explication systémique non plus au fait que la sous-population «Langues anciennes» affiche les compétences les plus équilibrées et les plus homogènes (ce qui en fait la sous-population où l'aptitude aux études supérieures est la plus développée), alors

que les élèves qui ont choisi certaines OS introduites récemment («Musique», «Arts visuels», «Philosophie/Pédagogie/Psychologie») sont à la peine. La hiérarchie entre les OS n'est pas tout à fait la même si l'on considère les résultats obtenus aux examens de maturité, et surtout, les différences entre les sous-populations sont moins marquées qu'aux tests. La proportion de notes insuffisantes à l'examen écrit de mathématiques est frappante chez plusieurs sous-populations. Elle frise parfois 50% («Langues modernes», «Arts visuels», «Musique») ou dépasse même ce seuil («Philosophie/Pédagogie/Psychologie»). Dans toutes ces sous-populations (hormis «Musique»), la note générale en mathématiques est également insuffisante dans près de 30% des cas. Les élèves qui ont choisi «Physique et applications des mathématiques» sont de loin les plus performants au test des compétences transversales. En queue de peloton, on retrouve les élèves qui ont choisi des OS récemment introduites: «Arts visuels», «Philosophie/Pédagogie/Psychologie» et, dans un domaine, «Musique».

Les écarts de performances entre les sexes constatés dans EVAMAR II rejoignent dans une large mesure les résultats d'études antérieures (Halpern & LaMay, 2000; Köller & Klieme, 2000; Stanat & Kunter, 2001; Dresel, Stöger & Ziegler, 2006). Les bachelières s'en sortent plutôt mieux en langue première et les bacheliers en mathématiques et au TCT, où les sciences expérimentales jouent un grand rôle. Selon l'étude PISA de 2003, il n'existe qu'un domaine où un déficit marqué persiste entre les sexes dans l'ensemble des pays étudiés. Il s'agit de l'aptitude à la lecture, où les garçons sont nettement plus faibles ($d_{\max} = 0,49$). La différence d'aptitude aux sciences expérimentales et à la résolution de problèmes n'est plus significative. En mathématiques, les garçons gardent un avantage, mais il est maigre ($d_{\max} = 0,19$; voir Zimmer 2006). L'étude PISA réalisée en 2006 confirme ces observations: pas d'écarts significatifs entre les sexes dans les compétences globales en sciences expérimentales et dans la résolution de problèmes, grande différence dans l'aptitude à la lecture et léger avantage pour les garçons en mathématiques (OCDE 2007). Dans l'évaluation EVAMAR II, les garçons s'en sortent mieux que dans les études PISA. Cela tient au fait que les adolescentes sont sur-proportionnellement présentes dans les gymnases en Suisse, ce qui diminue le niveau d'aptitude moyen des bachelières en termes relatifs.

Le groupe de cantons où le taux de maturités est le plus bas (le tiers inférieur) affiche de meilleurs résultats que ceux où il est le plus élevé (tiers supérieur). Selon les cas, les écarts sont légers à relativement importants. Le plus souvent, les notes aux examens de maturité sont également meilleures dans les cantons où le taux de maturités est faible, mais la correspondance n'est que partielle. La question de l'égalité des chances dans l'accès aux études supérieures se pose.

8.4 Pratiques d'évaluation

Les notes attribuées aux examens ne correspondent qu'en partie aux résultats des tests d'EVAMAR II. Les gymnases ne sont toutefois pas en cause. Les tests et les examens de maturité n'évaluent qu'en partie les mêmes compétences. L'angle de visée est plus ciblé pour les tests. On peut supposer aussi que les notes aux examens sont attribuées en tenant compte au moins partiellement du niveau de chaque classe (orientation d'après la norme sociale). Mais il y a des convergences structurelles entre les notes et les résultats des tests. Implicitement, le corps enseignant semble donc tenir compte aussi dans une large mesure de «standards de qualité objectifs» communs à toutes les classes (orientation critères). Des «anomalies» difficilement explicables subsistent. Ainsi, les notes attribuées en classe durant la dernière année de cours (la «notation continue») en biologie sont plus élevées chez les élèves qui ont choisi l'option spécifique «Langues anciennes» et «Arts visuels» que chez ceux qui ont opté pour «Biologie et chimie». Il est frappant aussi que beaucoup d'examens écrits de mathématiques et de langue première sont sanctionnés par une note insuffisante. Mais les notations continues et les notes à l'oral sont bien meilleures, «poussant» ainsi vers le haut la moyenne des notes de maturité.

8.5 Perception des compétences transversales propres

En complément aux études réalisées par Notter & Arnold (2003 et 2006), les bachelières et bacheliers ont été questionnés sur la manière dont ils perçoivent leurs propres compétences transversales. Les élèves estiment avoir des compétences satisfaisantes dans la plupart des domaines. Ils décèlent toutefois certains déficits, en particulier dans la manière de planifier de leur temps de travail et dans leur capacité de concentration.

Les étudiantes et étudiants questionnés par Notter & Arnold (2006) étaient également particulièrement critiques quant à leur capacité de bien gérer leur temps (p. 31). Pourtant, ils admettaient qu'il s'agit d'une aptitude essentielle (p. 52). Cet avis est partagé par les représentants du corps professoral interrogés dans le cadre du sous-projet A4 d'EVAMAR II (voir chapitre II). Environ 15% des professeurs qui jugent cette aptitude importante décèlent des déficits en la matière chez les étudiantes et étudiants, ce qui correspond aux réponses données par les bachelières et bacheliers.

La capacité de concentration des élèves ne figurait pas parmi les compétences qu'il était initialement prévu de tester dans le cadre d'EVAMAR II. Elle a été intégrée dans le questionnaire complémentaire sur la base d'indications fournies spontanément par le corps professoral, lorsqu'on lui a demandé quelles compétences mériteraient d'être testées au-delà de celles que l'on trouve déjà chez Notter & Arnold (2006). Cela a aussi conduit à élargir le champ de l'étude à d'autres compétences: esprit d'analyse et de synthèse, persévérance, capacité d'adopter d'autres points de vue que le sien et capacité d'argumenter. Dans tous ces domaines, les bachelières et bacheliers se sont attribués de bonnes notes en moyenne. La capacité d'adopter d'autres points de vue que le sien et la capacité d'argumenter rappellent une aptitude testée par Notter & Arnold dans leur étude de 2006: la compétence sociale. Les étudiantes et étudiants interrogés à l'époque estimaient être bien armés sur ce plan. Cela rejoint les observations faites par le corps professoral dans le cadre d'EVAMAR II. Environ 97% des professeurs ne décèlent pas de déficits dans la capacité de travailler en équipe de leurs étudiantes et étudiants - compétence qu'ils jugent par ailleurs importante. Leur jugement est un peu plus sévère dans deux autres domaines: la gestion du temps et l'efficacité dans la préparation des examens. Mais la proportion de professeurs qui perçoivent des déficits en la matière reste modérée (8%). En moyenne, les bachelières et bacheliers estiment également n'avoir pas vraiment de problèmes dans ces domaines.

Divers aspect de la personnalité de l'élève peuvent jouer un rôle dans le succès ultérieur de ses études (voir point 1.2.1). EVAMAR II en a exploré deux: le souci d'être performant (items sur la persévérance) et la capacité d'autodiscipline (items intégrés dans l'échelle mesurant l'efficacité de la préparation des examens). Si l'on en juge par les réponses des bachelières et bacheliers, un certain optimisme est de mise. Les valeurs enregistrées pour ces items sont plutôt élevées, ce qui constitue une bonne base pour leur aptitude aux études. Toutefois, on peut souhaiter une amélioration de la capacité de concentration des élèves, celle-ci faisant partie d'une bonne autodiscipline.

La capacité de traiter des informations en profondeur (*Deep Processing*) favorise également l'aptitude aux études. Dans le cadre d'EVAMAR II, certains aspects de cette dimension ont été analysés en tant que facteurs de motivation, à travers l'indicateur «Esprit d'analyse et de synthèse». Les bachelières et les bacheliers estiment que ce n'est pas leur compétence la plus développée. Ils pensent être plus doués pour préparer leurs examens efficacement par exemple. Ils disent néanmoins qu'il leur arrive «parfois» à «souvent» d'établir des liens entre différentes matières. Les élèves éprouvent donc le besoin de vraiment comprendre ce qui leur a été enseigné et pas seulement d'apprendre par cœur ou de réussir superficiellement. Ce souci de maîtrise (*Mastery Goals*) est de bon augure pour leurs études futures.

V Sous-projet D1: analyse d'examens écrits de maturité

1 Généralités et vue d'ensemble

1.1 Objectifs et choix des disciplines

Le sous-projet D1 d'EVAMAR II était consacré à l'analyse qualitative des examens écrits de maturité des bacheliers et bacheliers de tous les cantons. Il s'agissait d'évaluer les exigences posées aux élèves sur le plan cognitif et sur celui du contenu, et d'examiner dans quelle mesure elles correspondent, d'une part, aux différentes composantes de l'aptitude aux études supérieures identifiées dans le cadre des sous-projets A et B et, d'autre part, aux tests menés dans le module C. Cette analyse a été effectuée en premier lieu pour les disciplines ayant fait l'objet des tests: langue première, mathématiques et biologie; on pourrait envisager de l'étendre à d'autres disciplines lors d'un nouveau projet.

1.2 Généralités sur les instruments d'analyse

Il existe plusieurs approches permettant d'analyser les exigences requises sur le plan cognitif lors d'examens. Le projet EVAMAR II a recouru principalement à deux types de modèles: celui d'Anderson et al. (2001), déjà employé dans le module A, et les modèles repris dans les modules B et C et développés en fonction des disciplines.

S'agissant de l'analyse des exigences cognitives, le modèle de Bloom et al. (1956 et 1974) s'est imposé, plus exactement la version soigneusement remaniée qu'en ont donnée Anderson et al. (2001) en tenant compte des derniers acquis de la psychologie cognitive (cf. tableau V.1). Les mérites de ce modèle sont de trois ordres: premièrement, il facilite les assertions comparatives entre les disciplines. Deuxièmement, il s'agit d'un des modèles pédagogiques et didactiques les plus appréciés et utilisés des dernières décennies³⁵, qui gagne encore en importance dans la discussion actuelle sur les grilles de compétences (cf. par ex. Klieme & Leutner, 2006, p. 8). Il propose la grille d'analyse de loin la plus solide et la mieux pensée pour les exigences cognitives de performances scolaires. Troisièmement, il tient également compte de modèles spécifiques pour l'analyse des contenus propres aux disciplines. Une deuxième vérification a eu pour but de s'assurer que le modèle général pour les processus cognitifs était de nature s'adapter à chaque discipline analysée.

³⁵ Le modèle original a été traduit déjà en 22 langues.

Tableau V.1: Tableau des catégories et sous-catégories des processus cognitifs selon Anderson et al. (2001) et des définitions adaptées dans le cadre d'EVAMAR II (appliquée pour l'instant seulement à la biologie).

Catégories et sous-catégories	Processus cognitifs	Autres désignations	Définitions adaptées pour EVAMAR II (appliquées pour l'instant à la biologie)
	<i>[traduction d'Anderson et al. (2001), p. 63–92]</i>	<i>[traduction d'Anderson et al. (2001), p. 63–92]</i>	<i>[d'après Anderson et al. (2001), p. 63–92]</i>
1	rappeler faire appel à un savoir pertinent de sa mémoire à long terme		
1.1	reconnaître	identifier	trouver dans sa mémoire à long terme un savoir en relation avec les données reçues
1.2	Faire appel	remémorer	faire appel à des connaissances pertinentes retenues par la mémoire à long terme
2	comprendre saisir des contenus d'enseignement et en dégager du sens (oralement, par écrit, par une représentation graphique)		
2.1	interpréter	éclaircir, expliquer, paraphraser, reformuler, présenter autrement, traduire, transposer, retravailler	transposer des données reçues en modifiant la forme de présentation
2.2	donner des exemples, illustrer, montrer de façon paradigmatique	clarifier, illustrer, représenter de façon imagée	trouver un exemple ou une situation imagée qui illustre une notion ou un principe
2.3	classifier	catégoriser, subsumer	décider de l'appartenance d'une instance à telle ou telle catégorie (de notions ou de principes)
2.4	résumer	abstraire, généraliser	<ul style="list-style-type: none"> à partir de données reçues, trouver un sujet général ou résumer les éléments principaux (avec un processus personnel d'intégration) étudier les données reçues pour en tirer la quintessence, la pointe, «l'enjeu» et résumer (avec un processus personnel d'intégration)
2.5	déduire	conclure, inférer, extrapoler, interpoler, prévoir, estimer	<p>à partir des données reçues, dégager des conclusions logiques, par exemple donné: série d'exemples ou d'instances ou: donné: une analogie: A est à B ce que C est à D</p> <p>trouver: en comparant les exemples ou les instances, reconnaître des relations, des concepts ou des principes, en d'autres termes:</p> <ul style="list-style-type: none"> identifier un modèle dégager la règle de synthèse -> extrapolation trouver le terme analogique D identifier l'élément étranger
2.6	comparer	contraster, opposer, ranger, rechercher des analogies	<ul style="list-style-type: none"> reconnaître les convergences et les différences entre des idées, des objets et autres ordonner et classer des idées, des objets et autres au sein d'un même niveau hiérarchique (anglais: mapping)

Catégories et sous-catégories	Processus cognitifs	Autres désignations	Définitions adaptées pour EVAMAR II (appliquées pour l’instant à la biologie)
2.7	expliquer	construire des modèles et des mécanismes	appliquer à des éléments que l’on découvre un modèle, appris auparavant, décrivant un rapport de cause à effet appliquer signifie: <ul style="list-style-type: none">• si le donné=le ou les effet(s): utiliser le modèle pour remonter aux causes• si le donné=la ou les cause(s): utiliser le modèle pour conclure aux effets
3	appliquer appliquer ou utiliser une procédure dans une situation donnée		
3.1	exécuter	–	domaine commun pour 3.1/3.2 appliquer une procédure (étapes qui s’enchaînent ou qui ont pu être apprises) à un problème plus ou moins connu
3.2	implémenter	appliquer, faire usage de	
4	analyser diviser une donnée reçue en ses éléments constitutifs, définir les rapports mutuels de ceux-ci et la relation qu’ils entretiennent avec la structure ou l’objectif d’ensemble		
4.1	différencier	distinguer, séparer, accentuer, sélectionner	discerner des éléments pertinents et d’autres qui ne le sont pas au sein des données reçues (sans processus personnel d’intégration)
4.2	organiser	trouver une logique interne, intégrer, esquisser les contours, analyser, structurer	<ul style="list-style-type: none">• déterminer comment des éléments prennent place et interagissent au sein d’une structure• structurer des données reçues et les transposer sous une autre forme (attribuer des sous-titres, construire un tableau, élaborer un diagramme en bâtons, un histogramme, un topogramme, etc.)
4.3	se référer à quelque chose	déconstruire	à partir des données reçues, reconnaître le point de vue d’un auteur, les distorsions de perception et de jugement, sa représentation des valeurs ou ses intentions
5	évaluer émettre un jugement sur la base de critères et de standards		
5.1	vérifier	mettre des éléments en relation, détecter des erreurs, vérifier, tester	détecter des illogismes ou des erreurs au sein d’une démarche ou d’un produit; décider si la démarche ou le produit est en soi logique; déterminer l’effectivité d’une procédure (telle qu’elle est appliquée)
5.2	examiner de façon critique	apprécier	déceler, en fonction de critères externes, les points faibles d’un processus ou d’un produit; définir la cohérence externe d’un produit; reconnaître si une procédure est appropriée à un problème donné
6	créer réunir des éléments pour former un ensemble fonctionnel cohérent; articuler des éléments selon une nouvelle structure		
6.1	construire développer	poser des hypothèses	élaborer des hypothèses
6.2	planifier	donner forme	élaborer un mode de faire pour résoudre un problème posé
6.3	produire	construire	inventer un nouveau produit

Les résultats qu'il est possible d'obtenir à l'aide de cette grille d'analyse sont significatifs jusqu'à un certain point. Classer des tâches d'examens dans des catégories prédéterminées comporte inévitablement une part d'incertitude, comme la recherche en pédagogie l'a reconnu depuis longtemps. Cazden résume le problème en ces termes: «Penser des questions en termes de niveau de difficulté cognitive est probablement heuristiquement utile pour les enseignants, mais fondamentalement imprécis pour la recherche»³⁶. La restriction la plus importante est la suivante: le niveau d'exigence d'un énoncé d'examen dépend aussi toujours de l'enseignement qui a précédé. Une question considérée comme exigeante au premier abord taxonomiquement parlant peut fort bien se révéler être une simple reproduction de connaissances si sa solution a pu être mémorisée durant la phase de préparation. Les examens de maturité «faits maison» peuvent donner lieu à ce type de démarche qui consiste à «enseigner pour les tests» ou à «faire des tests pour confirmer l'enseignement»³⁷. Le projet EVAMAR II n'avait pas les moyens de se référer à l'enseignement antérieur, et cette tâche ne faisait d'ailleurs pas partie du projet de mandat. D'ailleurs, Carlsen signale (1991) qu'il n'est pas pensable, en termes d'économie de la recherche, de produire des pièces pour attester sans lacune de toute la démarche d'enseignement jusqu'à l'examen. Une manière élégante de résoudre le problème est d'interroger les examinateurs qui ont rédigé les données d'épreuve. Toutefois, il s'avère que maints enseignants ne sont pas familiarisés à l'utilisation de schémas cognitifs comme ceux d'Anderson et al. (2001) et que par ailleurs ils n'affectionnent guère de classer les questions qu'ils posent dans la catégorie 1.2, suspecte de révéler un manque d'imagination et la tendance à surcharger inutilement les bacheliers et bachelères de connaissances de détail. A cela s'ajoute que si l'on avait voulu garantir la fiabilité des décisions de catégorisation, il aurait fallu donner aux enseignants une formation préalable approfondie en matière de taxonomie, ou – pour chaque énoncé – engager des entretiens entre l'enseignant et les collaborateurs d'EVAMAR II. Une étude récemment publiée dans la revue «Science» fonde le sérieux d'une démarche taxonomique uniquement sur des codeurs non informés de l'enseignement qui a précédé (Zheng et al., 2008). Les auteurs admettent aussi que les indices subtils donnés dans l'énoncé d'examen, combinés avec l'expérience de codeurs qui sont aussi au bénéfice d'une pratique de l'enseignement, amènent à attribuer les questions à la catégorie la plus significative. En outre, la classification taxonomique des énoncés de tâches promet des renseignements intéressants même si l'on ignore leur part de relatif dû à l'enseignement qui a précédé. Les exigences des tâches à accomplir, telles qu'elles se posent indépendamment de ce qui a pu être travaillé et entraîné dans les leçons, brossent le tableau de connaissances et de compétences réellement attendues, bref elles dessinent un «standard» auquel on mesure les bacheliers et bachelères. Certains experts peuvent faire appel à un savoir intériorisé pour résoudre nombre de tâches qui demandent aux novices des performances cognitives élevées; de même en situation d'examens, une même tâche – tant dans une discipline que pour des compétences transversales – peut exiger des efforts de réflexion fort différents d'un bachelier à l'autre, selon les expériences individuelles. Au bout du compte, les uns et les autres ont maîtrisé la situation concrète s'ils arrivent à la même solution correcte. L'analyse effectuée dans le sous-projet D1 caractérise les tâches proposées indépendamment de l'histoire des personnes. En d'autres termes, on ne peut déterminer que **le niveau d'exigences cognitives potentiel** – ce qui n'est déjà pas mal (cf. le projet COACTIV, Jordan et al., 2006). La même difficulté obère les tests standardisés, à un degré moindre pourtant: la même question d'un test peut constituer pour l'une des participants au test une simple opération de reproduction du savoir et requérir d'un autre des tâches de transfert assez complexes. Il reste cependant intéressant d'analyser taxonomiquement la difficulté potentielle des tâches sans se référer au vécu individuel. Cela dit, d'autres questions restent ouvertes, comme celle de la comparabilité du potentiel de tâches avec les solutions réellement attendues dans les examens et avec les résultats, en points et en notes, obtenus par les étudiants.

Outre l'analyse des exigences cognitives, EVAMAR II a recensé et catégorisé les contenus des données qui ont été présentées en 2007 aux candidats aux examens de maturité des cantons suisses. L'étude visait à obtenir, pour chaque discipline étudiée, un descriptif synthétique fondé sur les grilles de compétences adoptées. Dans les

³⁶ «Thinking about questions in terms of some scale of cognitive difficulty is probably still heuristically useful for teachers, but inherently imprecise for research» (p. 453).

³⁷ «Teaching to the test» ou «testing to the teaching».

disciplines biologie et mathématiques - ayant fait l'objet des tests - , on a structuré les énoncés sur la base des grilles de compétences déjà élaborées qu'il a fallu parfois compléter en y intégrant les catégories de contenus non prévus dans les tests du module C, mais qui faisaient partie de la matière d'examen dans les épreuves de maturité analysées.

2 Description de l'échantillon

L'échantillon retenu comporte les mêmes écoles et les mêmes classes que dans le module C (cf. chapitre IV, section 2). Les données demandées aux écoles nous sont parvenues presque intégralement. Les gymnases du Tessin - le Liceo cantonale di Bellinzona mis à part - ont aussi participé à ce sous-projet.

3 Résultats

3.1 Langue première

Pour la langue première, l'analyse a porté sur les données des épreuves d'examen d'un ensemble de 122 écoles (68 en Suisse alémanique [D], 50 en Suisse romande [F] et 4 au Tessin [I]), soit 314 classes concernées en tout (D 106, F 174, I 34). Les données de maturité de ces 314 classes représentaient 152 examens différents (D 94, F 54, I 4): plusieurs classes avaient donc des examens en commun. Les 314 examens de maturité (en partie communs) comportaient un total de 872 énoncés ou problèmes partiellement communs (D 446, F 264, I 162).

3.1.1 Les outils d'analyse

Il a fallu entreprendre des recherches étendues dans la littérature et consulter des experts avant de déterminer quel était l'instrument le plus approprié (et le plus largement accepté) pour analyser le potentiel d'exigences cognitives. Notons que les résultats sont modestes: s'agissant de la langue première en effet, depuis la taxonomie déjà ancienne de Bloom et al. (1956), aucune nouvelle démarche mieux à même de définir des exigences cognitives de données d'examens n'a réussi à s'imposer. Une récente adaptation anglaise de la taxonomie de Bloom par Narciss & Proske (2001) a été utilisée dans les pays germanophones. Des études suisses recourent déjà à ce modèle, qui est conçu avant tout pour le niveau universitaire, mais qui se révélera probablement susceptible d'être étendu au domaine de la langue première. A notre avis cependant, la refonte de Narciss ne présente pas les atouts qui auraient pu compenser le désavantage de sa diffusion encore limitée. Il n'y avait donc pas d'autre solution, pour la discipline langue première, que d'employer la taxonomie d'Anderson et al. revisitée (2001, cf. section 1.2, tableau V.1). Les travaux préparatoires en vue de l'analyse des exigences posées aux élèves sur le plan cognitif en langue première ont consisté avant tout en une transposition de la grille d'analyse - avec ses définitions et ses exemples de catégories et de sous-catégories qui formulés d'abord pour les sciences expérimentales - afin qu'elle puisse s'adapter aux exigences et aux spécificités des performances langagières.

L'analyse des contenus d'examens en langue première a pris en compte les paramètres suivants):

- la forme de composition demandée (essai, dissertation, explication de texte et autre);
- le type du texte de départ figurant sur la donnée soumise aux élèves (poème, récit, texte documentaire ou autre);
- le siècle auquel appartient le texte soumis à l'examen;
- le poète ou le prosateur auteur du texte, son origine et son sexe;
- le ou les thèmes de la donnée ou du texte de départ;

On a relevé aussi les éléments purement quantitatifs suivants:

- la longueur des textes de départ soumis aux élèves;
- le nombre des documents / textes de départ et des possibilités de choix.

On a également saisi d'autres paramètres des examens:

- le type et le nombre de notes explicatives ajoutées par l'enseignant (lexique, contexte, etc.);
- le genre et le nombre des moyens auxiliaires autorisés (dictionnaire, ordinateur, etc.);
- la durée de l'examen (quand elle était indiquée).

Les pages qui suivent présentent les résultats de l'analyse effectuée dans l'ordre des éléments cités.

3.1.2 Les contenus des examens en langue première

3.1.2.1 La forme de l'épreuve

En règle générale, il règne une unité de doctrine dans toute la Suisse quant au type d'épreuve soumis aux bacheliers et bacheliers pour les examens écrits en langue première: à chaque fois, on demande à l'élève de produire une composition écrite de plusieurs pages. Nous n'avons vu aucune donnée de langue première – ni en allemand, ni en français, ni en italien – qui aurait soumis des problèmes de nature purement grammaticale ou posé des questions directes de compréhension ou de connaissances portant sur un domaine partiel de la littérature, de la théorie ou de l'histoire littéraire, ou encore de l'histoire de la culture ou de la langue. Dans tous les cas, les bacheliers et bacheliers suisses doivent créer, dans le cadre horaire de quatre heures en principe, un produit élaboré de façon personnelle et présenté en copie manuscrite (rarement avec l'aide d'un ordinateur portable personnel ou mis à disposition par l'école). Cependant, cette harmonie de surface dissimule un champ très varié de différences qu'il est intéressant d'étudier. Celles-ci figurent dans le tableau V.2 pour ce qui est de la Suisse alémanique.

Tableau V.2: Les formes de composition à l'examen de maturité dans les gymnases suisses alémaniques

Formes de composition Deutsch /Allemand	Titre 1	Option 2	Option 3	Option 4	Total (N)	Groupes (N)	Fréquence (%)	Groupes (%)
Interpretation / Explication de texte	54	2	0	0	56		10.6%	
Vergleichende Interpretation / Comparaison de textes	14	0	0	0	14		2.6%	
Analyse	13	0	0	0	13	83	2.5%	16%
Erörterung (freie) Dissertation libre	73	6	0	0	79		15.0%	
Erörterung (literarische) / Dissertation (sujet littéraire)	18	4	0	0	22		4.2%	
Erörterung (textgebundene) / Dissertation (sujet général)	19	0	0	0	19		3.6%	
Argumentation / Texte argumentatif	6	2	0	0	8	128	1.5%	24%
Essay / Essai	25	16	3	0	44		8.3%	
Betrachtung / Réflexion	5	0	0	0	5	49	0.9%	9%
Kurzgeschichte, Erzählung / Nouvelle, narration	21	8	0	1	30		5.7%	
Glosse / Commentaire	0	2	0	0	2		0.4%	
Satire / Satire	2	0	1	0	3		0.6%	
Groteske / Trait grotesque	1	0	0	0	1		0.2%	
Märchen, Parabel / Conte, parabole	0	4	0	0	4		0.8%	
Gedicht / Poème	0	1	0	2	3		0.6%	
Szene / Scène	0	0	3	0	3	46	0.6%	9%
Rede / Discours	5	1	1	1	8		1.5%	
Brief / Lettre	0	7	0	0	7		1.3%	
Zeitungsartikel / Article de journal	2	1	0	0	3		0.6%	
Dialog / Dialogue	0	0	4	1	5		0.9%	
Tagebuch / Journal intime	0	0	0	4	4	27	0.7%	5%
freie Aufsatzform Forme libre	188	4	3	0	195	195	36.9%	37%
Total «Formes de composition»	446	58	15	9	528		100.0%	100%
Dont: forme imposée	133				133		29.8%	

Les enseignants suisses alémaniques sont bien plus de la moitié (D 63%), à donner des consignes formelles précises quant au «produit fini» qu'ils attendent: dans un quart des cas environ, la forme prescrite est celle du texte discursif (dissertation, texte argumentatif), pour un sixième des compositions, on demande une explication ou une analyse de texte, 10% des enseignants attendent une réflexion écrite élevée sous la forme de l'essai, et 10% également autorisent aussi les formes délicates du récit. Quelques (très rares) enseignants proposent à leurs élèves de rédiger un article fictif de journal, un discours ou une lettre. Si d'un côté la plupart des données laissent le choix du genre de composition (avec deux, trois ou même quatre variantes), de l'autre, dans un quart des compositions, la forme du texte était strictement imposée. Un taux de 37% des bacheliers et bacheliers de langue allemande peuvent choisir en toute indépendance et sans restriction ni suggestions la forme qu'ils veulent donner à leur composition (c'est la «composition libre»): l'enseignant n'explicite alors

aucune consigne d'ordre formel, et le candidat est chargé de faire le choix de la forme appropriée en fonction du type de donnée ou des textes proposés.

Le tableau V.3 recense les formes de composition trouvées en Suisse italienne. La forme de la composition libre (*componimento*) s'y trouve encore plus fréquemment représentée (48%). Les enseignants ont donné une consigne formelle claire dans 52% des cas, la plupart du temps pour demander une analyse (40%), d'autres fois (12%) pour inviter à d'autres possibilités (*riflessione, confronto di testi*). Aucune des données de Suisse italienne que nous avons vues ne permettait un choix entre plusieurs formes de composition.

Tableau V.3: Les formes de composition à l'examen de maturité en Suisse italienne

Formes de composition en italien	Titre 1	Fréquence (%)
Componimento / Composition libre	78	48%
Analisi (poetico) / Analyse de texte poétique	34	21%
Analisi (narrativo) / Analyse de texte narratif	31	19%
Confronto di testi / Comparaison de textes	11	7%
Riflessione / Réflexion	8	5%
Total «Formes de composition»	162	100%
Dont: forme imposée	162	100%

Pour ce qui est de la Suisse romande (tableau V.4), la revue des formes de composition requises se présente ainsi: dans à peu près un cas sur deux, la forme attendue est celle de la «dissertation», appliquée soit à des sujets littéraires, soit à des sujets généraux et qui correspond au terme allemand «Erörterung». On recense un total de douze 12 formes différentes de composition. Outre le «commentaire composé (littéraire)» (env. 9%), l'équivalent de l'«Interpretation» des Alémaniques, et du texte argumentatif (pratiquement 2%), seul sort du lot, avec 5%, le genre appelé «évocation». Il s'agit de ce que l'on nommerait en allemand «Erweckung», un texte appelé à présenter en une langue imagée et soignée les sentiments, impressions et images suscités par des textes poétiques. Ce genre d'épreuve écrite n'apparaît pas sous cette forme dans les deux autres régions linguistiques et constitue une spécificité des cantons de Suisse romande.

Tableau V.4: Les formes de composition en Suisse romande

Type d'exercice	Formes de composition correspondantes en allemand	Sujet 1	Option 2	Fréquence
Etude de texte	Analyse	1		0.4%
Commentaire composé (littéraire)	Interpretation (immer literarisch)	24		9.1%
Comparaison de textes	Vergleichende Interpretation	6		2.3%
Dissertation Sujet littéraire	Erörterung literarische	71		26.9%
Dissertation libre	Erörterung (freie)	0		0.0%
Dissertation Sujet général	Erörterung (textgebundene)	62		23.5%
Texte argumentatif	Argumentation	6		2.3%
Essai	Essay	0		0.0%
Narration	Kurzgeschichte, Erzählung	1		0.4%
Commentaire de texte		5	1	1.89%
Evocation		13	15	4.9%
Méditation libre	freie Aufsatzform	0	1	0.0%
Texte libre	freie Aufsatzform	1	1	0.4%
Sans indication	freie Aufsatzform	74		28.0%
Total «Formes de composition»		264	18	100.0%
Dont: forme imposée			246	93.2%

Pour conclure, la situation se présente très différemment dans les trois régions linguistiques. Tandis que la Suisse italienne connaît cinq formes de composition aux examens et que souvent le texte tout à fait libre est aussi autorisé, on trouve douze formes de textes en Suisse romande et 22 en Suisse alémanique. Les gymnases de Suisse romande prescrivent en général une forme déterminée (72%) et la plupart du temps, comme au Tessin (I 100%), aucune variante n'est possible (F 93%). A l'inverse, les bachelières et bacheliers de Suisse alémanique peuvent fréquemment choisir (dans 313 cas sur 446 données examinées, soit 70%) entre diverses formes de composition suggérées ou même prendre une décision personnelle en l'absence de toute consigne.

3.1.2.2 Les textes de départ soumis aux élèves

Par texte de départ, on désigne ici les courtes citations ou extraits de textes d'une certaine longueur qui précèdent la consigne proprement dite ou lui font suite, parfois même jointes sur feuilles séparées. Il incombe à chaque candidat et candidate de prendre en compte ces documents et de s'y référer dans sa composition.

La plus grande part des textes de départ sont issus du domaine des «textes littéraires» (D 41%, F 62%, I 57%). Les textes documentaires sont nettement minoritaires (D 20%, F 21.7%, I 9%) et d'autres documents ressortissent des «textes philosophiques» (D 12%, F 14%, I 28%). Très rarement, les documents sont des photos, dessins ou caricatures (D 4%, F 1%, I 0%); au Tessin, on joint parfois des illustrations complémentaires, notamment les portraits des auteurs (I 14%). En vertu d'une habitude régionale, une part non négligeable des documents soumis en Suisse alémanique, soit 16% en tout, sont des textes personnels de l'enseignant ou l'enseignante, par exemple des survols récapitulatifs, des commentaires de l'actualité ou de la marche du temps, ou des questions de morales et des réflexions d'éthique à propos de phénomènes de société (cette variante en F: 2%, I 6%). (Pour un reliquat de 6% de données de Suisse alémanique, le type de texte n'a pu être clairement attribué à aucune catégorie.)

Les tableaux V.5, V.6 et V.7 présentent la statistique détaillée sur ce point pour la Suisse alémanique, la Suisse romande et le Tessin respectivement.

Tableau V.5: Les types de textes soumis aux élèves dans les épreuves de maturité de Suisse alémanique

Textes de départ «allemand»: types de textes / d'extraits		Nombre (N)	Fréquence (%)	Groupe (%)
Textes littéraires	aphorisme	53	26%	
	anecdote, parabole	4	2%	
	poème	70	34%	
	récit, nouvelle	40	20%	
	roman	8	4%	
	pièce de théâtre	13	6%	
	correspondance	8	4%	
	autobiographie / interview	7	3%	
Total «Textes littéraires»		203	100%	41%
Textes philosophiques	aphorisme	45	76%	
	traité	14	24%	
Total «Textes philosophiques»		59	100%	12%
Autres textes	lettre de lecteur	3	3%	
	exposé / discours	7	7%	
	essai	25	25%	
	article de journal	37	37%	
	texte publicitaire	6	6%	
	monographie	14	14%	
	monographie scientifique	4	4%	
	site Internet	4	4%	
Total «Autres textes»		100	100%	20%
Textes de l'enseignant	questions	22	28%	
	commentaires	57	72%	
Total «Textes de l'enseignant»		79	100%	16%
Images	photographie	6		
	reproduction de tableau	2		
	dessin / graphique	1		
	caricature / bande dessinée / dessin humoristique	10		
Total des images	images sans titres / titre	21		4%
Type de texte non identifié		30		6%
Total «Textes» (tous les types)		492		100%

Les tableaux présentent aussi la répartition de détail au sein des catégories générales de textes: on remarque ainsi qu'en Suisse alémanique plus du tiers des textes littéraires soumis aux candidats sont des poèmes, alors que 6% seulement sont empruntés à des œuvres dramatiques; pour les autres textes (ou textes documentaires), 37% d'entre eux sont des articles de journaux, et seulement 4% ont été clairement repris d'un site Internet.

En Suisse romande, plus du tiers des documents de départ littéraires sont tirés de romans, tandis que pour les poèmes et les essais, ce taux est d'à peine un quart pour chaque genre. Des articles de journaux ne sont employés que dans 13% des cas.

Tableau V.6: Les types de textes soumis aux élèves dans les épreuves de maturité de Suisse romande

Textes de départ «français»: types de textes / d'extraits		Nombre (N)	Fréquence (%)	Groupe (%)
Textes littéraires	roman	18	9.7%	
	aphorisme	42	22.6%	
	poème	11	5.9%	
	récit	67	36.0%	
	drame	5	2.7%	
	essai	42	22.6%	
	correspondance	1	0.5%	
	autobiographie	0	0.0%	
Total «Textes littéraires»		186	100.0%	62.0%
Textes philosophiques	aphorisme	20	46.5%	
	traité	23	53.5%	
Total «Textes philosophiques»		43	100.0%	14.3%
Autres textes	lettre de lecteur/lectrice	0	0.0%	
	exposé/discours	0	0.0%	
	essai	56	86.2%	
	ouvrage spécialisé	0	0.0%	
	ouvrage de vulgarisation	0	0.0%	
	article de journal	9	13.8%	
	publicité	0	0.0%	
Total «Autres textes»		65	100.0%	21.7%
question enseignant		6		
commentaire enseignant		0		
Total «Textes de l'enseignant»		6		2.0%
Total «Textes» (tous les types)		300		100.0%

Au Tessin, l'accent est mis sur les poèmes (20%) et les récits (13%), ou sur les textes philosophiques (28%).

Tableau V.7: Les types de textes soumis aux élèves dans les épreuves de maturité de Suisse italienne (Tessin)

Textes de départ «italien»: types de textes / d'extraits		Nombre (N)	Fréquence (%)	Groupe (%)
Textes littéraires				
	aphorisme	0	0%	
	poème	34	20%	
	récit	23	13%	
	roman (2 x Manzoni)	11	6%	
	nouvelle	8	5%	
	pièce de théâtre	0	0%	
	correspondance	0	0%	
	autobiographie	11	6%	
	article sur un roman	11	6%	
Total «Textes littéraires»		98	57%	57%
Textes philosophiques				
	aphorisme	7	4%	
	traité (discussion), Argumentation	15	9%	
	réflexion	16	9%	
	discours	11	6%	
Total «Textes philosophiques»		49	28%	28%
Autres textes				
	lettre de lecteur	0	0%	
	exposé / discours	0	0%	
	essai	8	5%	
	monographie	0	0%	
	monographie scientifique (texte historique)	7	4%	
	article de journal	0	0%	
	texte publicitaire	0	0%	
Total «Autres textes»		15	9%	9%
Textes de l'enseignant				
	questions	0	0%	
	commentaires	11	6%	
Total «Textes de l'enseignant»		11	6%	6%
Type de texte non identifié		0	0%	
Total «Textes» (tous les types)		173		100%

3.1.2.3 Datation des textes

Tous les examens analysés dans notre pays font apparaître une tendance très marquée vers des textes récents ou très récents: la plupart des textes sont issus du 20^e siècle (D 37%, F 65%, I 39%), ou du 21^e siècle qui commence (D 42%, F 14%, I 9%). Il est rare que l'épreuve écrite de maturité mette les bacheliers et bacheliers en présence d'extraits de la littérature des 17 au 19^{es} siècles (D 14%, F 15%, I 27%). L'exception, au Tessin, est la proportion d'extraits de textes des 14 au 16^{es} siècles (20%), nettement plus élevée que pour les documents du 21^e siècle (9%).

Tableau V.8: Indications des siècles auxquels appartiennent les textes soumis aux élèves (épreuves de maturité en Suisse, trois régions linguistiques)

Siècle	D (N)	D (%)	D (%) Total	F (N)	F (%)	F (%) Total	I (N)	I (%)	I (%) Total
Antiquité / av. J.-C.	7	1,5%		0	0.0%		8	4.6%	
1 ^{er} – 13 ^e siècles	2	0.4%		0	0.0%		0	0.0%	
14 ^e – 16 ^e siècles	2	0.4%	0.8%	9	3.0%	3.0%	34	19.7%	24.3%
17 ^e siècle	8	1.7%		6	2.0%		14	8.1%	
18 ^e siècle	26	5.4%		4	1.3%		11	6.4%	
19 ^e siècle	34	7.1%	14.1%	36	11.8%	15.1%	22	12.7%	27.2%
Début du 20 ^e siècle	83	17.3%		73	23.9%		23	13.3%	
Fin du 20 ^e siècle	96	20.0%	37.2%	126	41.3%	65.2%	45	26.0%	39.3%
21 ^e siècle	121	25.2%		44	14.4%		16	9.2%	
21 ^e siècle (enseignant)	81	16.8%	42.0%			14.4%	0	0.0%	9.2%
aucune indication	21	4.4%		7	2.3%		0	0.0%	
Total recensé	481	100.0%		305	100.0%		173	100.0%	

3.1.2.4 Auteurs (hommes et femmes)

Parallèle à la tendance «moderniste» que l'on vient de commenter, la longue liste des poètes, prosateurs, dramaturges, essayistes cités par les enseignants dans les sujets d'examens témoigne de la diversité des intérêts: de l'ancien canon classique de l'enseignement de la littérature, seuls se distinguent, avec plus de cinq mentions pour chacun, les noms de Goethe, Brecht, Kafka, Fried, Rilke et Dürrenmatt; ils réunissent ensemble à peine un quart de tous les extraits littéraires (pour le détail, cf. tableau V.9).

Tableau V.9: Auteurs cités dans les épreuves de maturité de Suisse alémanique (extrait)

Auteurs des textes de départ soumis aux élèves «allemand»	Nombre (N)	Proportion (%)	Littéraire (%)
Enseignant	77	15.7%	
Journaliste	36	7.3%	
Auteur inconnu (dicton, sagesse populaire etc.)	15	3.0%	
Aucun texte de départ soumis	45	9.1%	
Aucune indication donnée	19	3.9%	
Ausländer, Rose	2	0.4%	1.0%
Bachmann, Ingeborg	2	0.4%	1.0%
Benn, Gottfried	5	1.0%	2.5%
Böll, Heinrich	1	0.2%	0.5%
Brecht	11	2.2%	5.4%
Domin, Hilde	2	0.4%	1.0%
Dürrenmatt	7	1.4%	3.4%
Eichendorff	2	0.4%	1.0%
Fried, Erich	7	1.4%	3.4%
Goethe	11	2.2%	5.4%
Hahn, Ulla	5	1.0%	2.5%
Handke	4	0.8%	2.0%
Hofmannsthal	2	0.4%	1.0%
Kästner, Erich	3	0.6%	1.5%
Kafka	8	1.6%	3.9%
Kaschnitz, Marie Luise	3	0.6%	1.5%

Auteurs des textes de départ soumis aux élèves «allemand»	Nombre (N)	Proportion (%)	Littéraire (%)
Kleist	2	0.4%	1.0%
Kunert, Günther	4	0.8%	2.0%
Lessing	3	0.6%	1.5%
Mann, Thomas	2	0.4%	1.0%
Nietzsche	3	0.6%	1.5%
Novalis	3	0.6%	1.5%
Rilke	6	1.2%	3.0%
Schiller	3	0.6%	1.5%
Tucholsky	3	0.6%	1.5%
Biermann	2	0.4%	1.0%
Grönemeyer	1	0.2%	0.5%
Blocher (dans sa fonction de conseiller fédéral)	3	0.6%	
Conseiller fédéral Couchepin	1	0.2%	
Conseiller fédéral Leuenberger	1	0.2%	
Total «Auteur des textes» (extrait)	304	59.6%	52.7%
Total des auteurs pour l'allemand	492	100.0%	

Il est un point où les enseignants de Suisse alémanique ne se démarquent pas de la pensée traditionnelle de la germanistique: la plupart du temps, ils citent dans leurs données des textes de littérature allemande originale, donc d'auteurs du domaine germanophone (voir tableau V.10). De cette proportion de 82% de textes de départ écrits en allemand, quelque 14% proviennent de la littérature et de la presse suisse. Société globalisée ou non, les textes traduits restent minoritaires dans les examens de langue première «allemand»: un nombre très restreint d'extraits appartiennent aux espaces francophone, asiatique ou anglophone.

Tableau V.10: Origine/nationalité des auteurs cités (épreuves de maturité d'allemand de Suisse alémanique)

Pays ou culture d'origine	Nombre (N)	Fréquence (%)	Groupe (%)
Suisse	69	14.0%	
probablement Suisse (enseignant / journaliste)	100	20.3%	
Allemagne	185	37.6%	
Autriche	36	7.3%	
Tchéquie (germanophone)	13	2.6%	81.9%
France	13	2.6%	
Italie	7	1.4%	
Anglophonie	21	4.3%	
Littérature hispanique	3	0.6%	
Littérature russe	4	0.8%	
Asie	9	1.8%	
autre	7	1.4%	13.0%
Non identifié	25	5.1%	5.1%
Total des auteurs pour l'allemand	492	100.0%	100.0%

Une relation comparable est observable en Suisse Romande (cf. tableau V.11): les citations directes d'auteurs de la littérature francophone sont en large majorité (84%), les citations traduites sont rares. Les auteurs de Suisse romande forment ensemble un peloton de 10%, et 3 de 30 documents sont probablement des traductions.

Tableau V.11: Origine/nationalité des auteurs cités (épreuves de maturité de français de Suisse romande)

Nationalité de l'auteur	Nombre (N)	Fréquence (%)	Groupe (%)
Suisse	24	7.9%	
Suisse, probablement (enseignant / journaliste)	6	2.0%	9.9%
Allemagne	5	1.7%	
Autriche	7	2.3%	4.0%
France /pays de l'est européen	36	11.9%	
France	188	62.3%	84.1%
Italie	5	1.7%	
Pays anglophone	2	0.7%	
Pays hispanophone	0	0.0%	
Pays russophone	0	0.0%	
Asie	0	0.0%	
autre	4	1.3%	3.6%
inconnu	25	8.3%	8.3%
Total des auteurs pour le français	302	100.0%	100.0%

En Suisse italienne, 17 des 20 textes cités³⁸ proviennent d'auteurs (hommes ou femmes) d'Italie; un écrivain appartient à la francophonie, un au monde anglophone et un à la Grèce antique (Aristote). Au Tessin, aucune donnée d'examen ne comporte de texte d'un compatriote.

On conclut donc que les trois régions linguistiques du pays recourent essentiellement à des textes littéraires en version originale (D 82%, F 84%, I 85%). Des citations traduites de langues voisines demeurent l'exception. Une référence marquée à la littérature d'auteurs suisses est observable chez les enseignants alémaniques principalement (14%), suivis des Romands (9.9%).

Les enseignants suisses demeurent très traditionnels quant à la manière dont ils ne tiennent pas compte des sexes pour choisir leurs extraits (cf. tableau V.12); dans chacune des trois régions linguistiques, les textes écrits par des auteurs féminins sont peu nombreux à avoir eu l'honneur d'une citation des examinateurs: les proportions d'auteurs femmes sont de 9% pour la Suisse alémanique, de 10% en Suisse romande et de 5% au Tessin. A cet égard, les exigences d'un enseignement équilibré qui intègre le genre font ressortir un net besoin de rattrapage et d'information chez les enseignants des trois régions linguistiques.

Tableau V.12: Genre (sexe) des auteurs cités (données des épreuves de maturité de toute la Suisse)

Genre des auteurs	D (N)	D (%)	F (N)	F (%)	I (N)	I (%)
féminin	40	8.7%	31	10.2%	8	5%
masculin	322	69.7%	266	87.2%	165	95%
Non identifié (enseignant / journaliste)	100	21.6%	8	2.6%	0	0%
Total recensé	462	100.0%	305	100.0%	173	100%

3.1.2.5 Thématique

Des 446 sujets de compositions possibles proposés dans les examens de maturité de Suisse alémanique sur lesquels a porté l'analyse, on a pu dresser une liste de 300 mots-clés. Ceux-ci ont été regroupés, parallèlement aux genres de textes, en thèmes littéraires, thèmes philosophiques et thèmes documentaires. Nous pouvons souligner

³⁸ Pour le Tessin, le compte se fonde sur l'ensemble des examens (I 4), lesquels, dans chacune des quatre écoles, sont organisés de façon commune pour toutes les classes.

plusieurs évidences à cet égard: d'abord la présence de sujets classiques liés à la littérature, tels l'amour, le désir, la crise, la guerre, la nature, la ville, le destin, le hasard, le caractère et les voyages etc. Ensuite, les fréquences notables de certains thèmes de métaréflexion comme «fonction de la littérature / effet des livres / la lecture comme acte d'aventure» ou «la langue et la réalité». Par contre, on est surpris de la rareté de problèmes ou de questions étroitement liés à la science de la littérature – par exemple sur la théorie des genres littéraires, sur la poésie engagée, le théâtre épique ou l'adaptation d'œuvres littéraires au cinéma. Parmi les sujets philosophiques se profilent – pour citer quelques exemples – des questions sur la responsabilité de l'homme, sur son immaturité (dont il est fautif), sur le courage, le bonheur et la compassion, sur la patrie et l'étranger, sur le passé et l'avenir, sur le rôle de la formation, de l'ordre et du bon sens populaire. Dans la plupart des cas, les sujets documentaires sont actuels et très variés: des centrales nucléaires et du rapport sur le réchauffement climatique à la fonction de l'Etat en passant par la religion, l'éducation, les salaires minimaux, les performances de pointe, l'assistance au décès, la démocratie, et on en passe. Un champ important est celui de la thématique proche des jeunes: la fascination de l'informatique, le monde virtuel, le téléphone cellulaire, l'ivresse et l'alcool, la criminalité juvénile, les anglicismes dans la langue courante, le rêve de devenir modèle, l'endettement, le droit de vote et d'éligibilité à 16 ans, etc.

En résumé, le champ thématique couvert par les compositions d'examens de maturité en Suisse alémanique témoigne tout à la fois d'exigence, de variété et d'ouverture à l'actualité. En composant les textes de leurs examens finaux, les bacheliers et bacheliers se confrontent à des questions de philosophie existentielle, à des problèmes éthiques pertinents pour le monde dans lequel ils vivent et vivront, ils réfléchissent ainsi tant à des phénomènes actuels qu'à des aspects éternels de l'humaine condition.

3.1.2.6 Longueur des textes soumis

Outre le contenu, la longueur (ou le nombre) des textes soumis aux élèves est un autre paramètre qui peut influencer sur le niveau d'exigence d'un examen. En effet, le travail d'un candidat n'est pas le même s'il doit tenir compte de plusieurs pages de dialogues d'une pièce de théâtre, s'il n'a qu'à prendre connaissance d'un aphorisme de deux lignes ou s'il est appelé à comparer un poème de 40 vers avec le contenu d'une nouvelle d'une centaine de lignes.

Sur ce point aussi, l'analyse a révélé une situation étonnamment hétérogène, tant entre les régions linguistiques qu'au sein même de celles-ci (cf. tableau V.13). Ainsi, on dénombre tout de même un cinquième de tous les examens de Suisse alémanique où les enseignants confrontent leurs bacheliers et bacheliers à des textes longs (jusqu'à 100 lignes) ou très longs (plus de 100 lignes). A l'autre extrême, une moitié des gymnasiens alémaniques se voient proposer des textes courts, c'est-à-dire de une à cinq lignes. Un quart seulement des examens alémaniques travaillent sur des textes de longueur moyenne (de 6 à 20 lignes). En Suisse romande, deux tiers des bacheliers et bacheliers reçoivent des textes courts comptant de une à cinq lignes, un huitième, des textes de longueur moyenne, et 8% seulement doivent se mesurer à des textes longs ou très longs.

Tableau V.13: Longueur des textes de départ soumis (épreuves écrites de maturité, les trois régions linguistiques)

Texte de départ	Longueur	D (N)	D (%)	Groupe D (%)	F (N)	F (%)	Groupe F (%)	I (N)	I (%)	Groupe I (%)
aucun	un mot-clé / un titre	45	9%		0	0%		0	0%	
court	1–2 lignes	129	26%		144	47.2%		18	10%	
	3–5 lignes	96	20%	46%	59	19.3%	66.5%	11	6%	17%
moyen	6–10 lignes	73	15%		13	4.3%		18	10%	
	11–20 lignes	66	13%	28%	25	8.2%	12.5%	50	29%	39%
long	21–40 lignes	48	10%		8	2.6%		23	13%	
	Plus 40 lignes	42	8.5%	18.5%	13	4.3%	6.9%	30	17%	30.5%
très long	plus de 100 lignes	10	2%		5	1.7%		23	13%	
	Œuvre entière	7	1.5%	3.5%	34	11.2%	12.8%	0	0%	13.5%
illustration	une photo, sans texte	21	4%	4%	3	1%	1%	0	0%	
	texte manque	0	0%		1	0.3%		0	0%	
Total textes (longueur)		492	100%		305	100%		173	100%	

On constate une plus grande unité dans la manière dont les enseignants procèdent quant au nombre de textes soumis aux candidats (cf. tableau V.14). La majorité des examens ne présentent qu'un seul texte de départ (D 75%, F 87%, I 93%), un huitième environ des énoncés fournis comportent deux textes (D 12%, F 13%, I 7%), et seuls 3% des examens de tous les examens de Suisse alémanique reposent sur trois textes ou plus, que l'on demande de comparer.

Tableau V.14: Nombre de textes de départ soumis (épreuves écrites de maturité, les trois régions linguistiques)

Nombre de textes / de photos soumis	D (N)	D (%)	F (N)	F (%)	I (N)	I (%)
Aucun texte / aucune photo	45	10%	0	0%	0	0%
Un texte (photo)	333	75%	233	86.6%	151	93%
Deux textes (photos)	53	12%	36	13.4%	11	7%
Trois textes (photos)	11	2%	0	0%	0	0%
Quatre textes (photos)	0	0%	0	0%	0	0%
Cinq textes (photos)	4	1%	0	0%	0	0%
Total (données d'examens)	446	100%	269	100%	162	100%
Total «Nombre de textes soumis»	492		305		173	

Seule particularité frappante: 10% des données d'examens de Suisse alémanique ne comportent aucun texte; dans ces cas-là, la tâche consiste à rédiger une composition à partir d'un titre constitué d'un simple mot-clé (par ex. «Apparence», «A l'écart», «En secret», «Limites», «Seuil»).

Il est donc possible qu'un candidat aux examens de maturité dans notre pays reçoive une unique feuille A4 comportant trois sujets avec les textes courts y afférents et puisse ainsi vite faire son choix pour se mettre d'emblée à la tâche, donc à la composition du texte. Mais il peut arriver aussi que l'élève commence l'examen devant un recueil de dix pages à compulser, comprenant un mélange de citations très brèves et d'annexes de plusieurs pages, et une gamme de sept sujets variés qui lui sont soumis: un tel cas de figure sollicite d'abord la compétence de lecture sélective à un haut degré, puisque le candidat doit définir lui-même le cadre de son travail et opérer un choix (cf. paragraphe suivant).

3.1.2.7 Possibilités de choix

Des 152 examens analysés, seuls quatre (dans quatre établissements et quatre cantons différents) ne présentaient que deux sujets à choix (cf. tableau V.15); situation opposée dans une classe à qui l'on a présenté neuf sujets de composition; dans sept cas, une sélection de huit sujets a été proposée (en Suisse romande). Tels sont les extrêmes pour le nombre des variantes. Une large majorité d'élèves (soit 71% des classes) a reçu un éventail de quatre ou cinq sujets; au Tessin, c'était le cas de toutes les classes examinées.

Tableau V.15: Nombre de sujets de composition proposés à choix (épreuves de maturité de Suisse alémanique et de Suisse romande)

Nombre de sujets à choix	Examens D (N)	Sujets D (N)	D (%)	Examens F (N)	Sujets F (N)	F (%)
Deux	3	6	3%	1	2	2%
Trois	19	57	20%	13	39	24%
Quatre	39	156	41%	16	64	30%
Cinq	26	130	28%	4	20	7%
Six	3	18	3%	8	48	15%
Sept	3	21	3%	5	35	9%
Huit	0	0	0%	7	56	13%
Neuf	1	9	1%	0	0	0%
Total	94	397	100%	54	264	100%

3.1.2.8 Explications complémentaires

On a observé une concordance élevée dans la manière dont les enseignants de Suisse alémanique et de Suisse romande font usage ou non de la possibilité de donner des indications auxiliaires sur la donnée d'examen. Dans la plupart des cas, aucune explication complémentaire n'est formulée (D 73,2%, F 74%). A l'inverse, le Tessin recourt assez souvent à ces précisions (47%), par exemple sous la forme de notes de bas de page (pour des explications de lexique, sur l'époque ou le contexte).

En Suisse alémanique (cf. tableau V.16), l'examen ne comporte que rarement de brèves informations de lexique (5,2%, par exemple pour les termes «éditorial», «assaut», (*Impetus*) «consensus», «nostalgie») ou des commentaires sur le contexte (7,3% par ex. pour «Réaction à une lettre de lecteur de la veille»), sur le genre du texte soumis ou sur son contenu (par ex. «Discours d'un conseiller fédéral sur l'action sociale»). Dans de rares cas, négligeables, l'enseignant fournit de longs renseignements sur le contenu (0,4%), sur la forme de composition demandée (par ex. les critères d'évaluation pour les textes créatifs sont l'originalité, l'imagination) ou sur la caractérisation de l'époque (1,3%)³⁹.

Des explications complémentaires concernant l'auteur ne sont pas très fréquentes non plus (D 31,5%, I 10%; F pas relevé); ces additions se limitent le plus souvent, en Suisse alémanique, aux dates de la vie de l'auteur, à la désignation de la profession ou à la nationalité (par ex. «politologue français, poétesse autrichienne, philosophe allemand»); dans quatre cas seulement (D 0,8%) figurent une brève autobiographie ou une photographie de l'auteur. A l'inverse, les examens de Suisse italienne joignent plus volontiers aux extraits soumis une biographie assez détaillée de l'auteur avec photo, dates, évolution et positionnement de ses œuvres dans son époque (10%).

Curieusement, seuls huit examens précisaient la longueur attendue pour la composition.

Tableau V.16: Explications complémentaires de l'enseignant (épreuves de maturité de Suisse alémanique)

Explications	D (N)	D (%)
Explications lexicales	28	5.2%
Contexte	39	7.3%
Epoque	7	1.3%
Brève introduction concernant le contenu	32	6.0%
Longue introduction concernant le contenu	2	0.4%
Sur le type d'épreuve	9	1.7%
Sur l'espèce de texte	27	5.0%
Aucune explication	393	73.2%
Total «Explications»	537	100.0%

3.1.2.9 Moyens auxiliaires

Un bon nombre des énoncés analysés ne disent rien des moyens auxiliaires autorisés (D 42%, F 74%); on admet que ce point a été éclairci avant les examens déjà. Pour les épreuves de Suisse alémanique qui contiennent des précisions explicites, la moitié des classes environ (53%) ont l'autorisation d'employer le Duden de l'orthographe (Duden für die deutsche Rechtschreibung) ou selon les cas aussi un autre dictionnaire (le Wahrig ou le Bertelsmann). Très rarement, deux dictionnaires sont permis, ou alors le Duden et l'aide électronique du programme informatique de correction automatique (4%). Il n'arrive presque jamais que tout moyen auxiliaire soit explicitement banni (2%). Un seul des examens prescrivait aux élèves de prendre à l'examen le texte d'une œuvre lue en classe. Il est intéressant de relever que certaines écoles comprenant plusieurs classes parallèles à l'examen ne règlent pas cette question de façon uniforme.

³⁹ Pour les travaux de Suisse romande et du Tessin, le relevé des indications n'a pas été aussi détaillé.

Par rapport à la Suisse alémanique, la Suisse romande montre en la matière une pratique différente, puisqu'un quart des candidats (24%) doivent venir à l'examen écrit munis des œuvres lues en classe: dans ces cas-là, l'enseignant attend manifestement des bacheliers et bacheliers qu'ils se réfèrent à l'œuvre concernée par le sujet traité et citent des extraits précis. Dans certaines classes, les résumés ou les notes personnelles sont aussi autorisés à l'examen (cf. tableau V.17).

Tableau V.17: Moyens auxiliaires autorisés (épreuves de maturité de Suisse romande, cumul de mentions possible)

Matériel autorisé	F (N)	F (%)
Dictionnaires	20	37%
Œuvres	13	24%
Résumé	5	9%
Notes	2	4%
sans indication	40	74%

Au Tessin, toutes les classes examinées peuvent utiliser deux dictionnaires: un dictionnaire usuel (Dizionario della lingua italiana) ainsi qu'un dictionnaire des synonymes (Dizionario dei sinonimi e contrari).

Un ordinateur est autorisé ou même prescrit dans seulement six des classes analysées en Suisse alémanique (5.7%). Au Tessin et en Suisse romande, nous n'avons trouvé aucune mention d'une utilisation possible de moyens informatiques.

3.1.2.10 Durée de l'examen

A peu près la moitié des données ne comportaient aucune indication quant à la durée de l'examen (D 49%, F 56%, I 0%). Dans l'autre moitié des cas, toutes les classes ou presque disposaient de quatre heures (D 50%, F 44%, I 100%), à l'exception d'une unique classe de Suisse alémanique qui a effectué son examen de langue première en trois heures seulement.

3.1.3 Exigences cognitives

Pour commencer, il est utile de préciser que toutes les compositions de l'examen de maturité en langue première correspondent à la catégorie principale 6, soit à l'exigence de «Créer» (réunir des éléments dans le but de former un ensemble fonctionnel cohérent; structurer des éléments selon un nouvel ordonnancement). Le genre de la composition engage tout particulièrement la performance cognitive de «créer un produit personnel». Cette dernière correspond, chez Anderson et al. (2001), à la catégorie 6.3.

En outre, chaque composition comporte une ou plusieurs performances supplémentaires en relation avec des niveaux d'exigences cognitives divers (selon Anderson et al., 2001). Celles-ci dépendent de la forme de composition attendue des bacheliers et bacheliers, du genre de texte soumis comme donnée initiale et de la formulation de l'énoncé. Chaque sujet proposé acquiert ainsi un profil qui lui est propre et qui peut concerner plusieurs catégories d'exigences cognitives, si l'enseignant a indiqué ses attentes en sériant plusieurs tâches ou phases successives. L'objet du codage, pour la langue première, a été la consigne explicite donnée, c'est-à-dire l'instruction du travail à faire telle qu'elle est formulée par l'enseignant avec des étapes le cas échéant. En revanche, on a renoncé à tenir compte des performances cognitives (partielles) implicites que les bacheliers et bacheliers devaient effectuer en plus pour maîtriser la tâche qu'on leur demandait (par exemple, lors de l'analyse d'un récit, le rappel de la mémoire à long terme de connaissances pertinentes sur l'auteur, l'époque ou les particularités du genre de texte).

Si aucune consigne n'était formulée en plus du sujet – ce qui, à la différence de la Suisse alémanique (8%), arrivait presque une fois sur deux au Tessin (41%) et plus fréquemment encore en Suisse romande (67%), le codage du niveau de l'exigence cognitive a été déduit du genre de texte de départ soumis et des thèmes contenus, ainsi que de la forme de composition attendue.

S'agissant de l'attribution des points aux sujets qui comprenaient plusieurs tâches partielles, on a utilisé une pondération adéquate pour en user équitablement par rapport aux sujets ne comportant qu'une tâche: si le thème comprenait quatre tâches partielles, on a appliqué un facteur de pondération de 0.25 à chacune des catégories d'exigences cognitives, pour un sujet avec deux tâches partielles, le facteur était de 0.5 à chacune; ce mode de calcul a permis de donner le même poids d'un point à chaque sujet.

Pour chaque niveau d'exigence le plus élevé, le codage s'est fait dans l'esprit d'un potentiel d'exigences (pour la justification, voir la section 1.2).

La double dénomination rend justice au fait déjà décrit que chaque composition résulte en principe d'un processus de création complexe (catégorie 6.3, Créer), mais qu'elle requiert aussi des prestations partielles qui sont fonction de l'accent de chaque sujet: la première partie de la double mention caractérise la catégorie de la création en soi, commune à toutes les compositions quel qu'en soit le thème. Le second verbe identifie la sous-catégorie dans laquelle on range la prestation partielle demandée explicitement selon le sujet. Il y a lieu de préciser que chaque composition de maturité englobe implicitement plusieurs catégories d'exigences cognitives et que la décision de codage des sous-catégories explicites doit se limiter, pour chaque énoncé, à considérer le potentiel d'exigences cognitives prépondérant.

Pour la Suisse alémanique, 446 énoncés, comprenant en tout 868 tâches partielles, ont fait l'objet d'un classement taxonomique. Si l'on considère la répartition d'ensemble effectuée sur les six niveaux principaux des exigences cognitives (cf. tableau V.18), on constate que presque un tiers des sujets d'épreuves écrites de maturité de Suisse alémanique sont attribués au niveau d'exigence «Créer –évaluer». Pour tous ces sujets, les bacheliers et bacheliers composaient un texte où ils devaient émettre un jugement sur la base de critères personnels et de «standards» à expliciter. Dans la catégorie dominante 5.2, il y a lieu d'analyser un phénomène ou une évolution de façon critique, de prendre position en utilisant une argumentation «pour et contre» et d'apprécier une thèse, une vision du monde ou une démarche. La plupart du temps, ces compositions avaient la forme d'une dissertation.

Tableau V.18: Analyse des exigences cognitives (épreuves de maturité de Suisse alémanique, catégories principales)

Catégories cognitives principales (avec pondération), allemand	Fréquence (D)
Créer – rappeler (1)	2.3%
Créer – comprendre (2)	19.1%
Créer – appliquer (3)	1.5%
Créer – analyser (4)	22.2%
Créer – évaluer (5)	32.0%
Créer – créer (6)	22.9%
Total des catégories cognitives (pondéré)	100.0%

Un peu moins du quart des sujets de composition relèvent du domaine Créer-analyser. C'est le cas des sujets qui demandent aux bacheliers et bacheliers d'analyser les textes soumis selon leurs éléments constitutifs et de définir les rapports entre les parties essentielles et la relation qu'elles entretiennent avec la structure d'ensemble. S'agissant de la catégorie 4.3 qui domine dans ce groupe, l'enjeu est de reconnaître le point de vue et les intentions de l'auteur et d'interpréter le texte. Fréquemment, les extraits à expliquer sont des poèmes ou des nouvelles. Un peu moins du quart aussi des thèmes de maturité ressortissent de Créer-crée. Ici, la tâche est en général complexe pour les candidats, qui, en l'absence de consignes précises et de texte de départ, doivent produire un texte complètement neuf, par exemple en inventant une nouvelle originale ou un composant un essai difficile. Un cinquième environ des sujets appartiennent au domaine Créer-comprendre. Dans ces cas-là, les bacheliers et bacheliers doivent être capables de dégager du sens des données qui leur sont soumises, c'est-à-dire, selon les cas, en trouvant des exemples personnels pour illustrer le propos (catégorie 2.2), en tirant des parallèles (catégorie 2.6) ou en élaborant un modèle «cause-effet» pour expliquer un phénomène (catégorie 2.7).

Il ressort que tant dans le domaine Créer-rappeler (où il s'agit de faire appel à des connaissances) que dans le niveau 3, où c'est l'application de savoirs qui est requise, les données d'examen sont souvent dépourvues de toute mention de la tâche à accomplir. On en déduit que la vérification de connaissances explicites ne fait pas partie des objectifs essentiels des épreuves écrites de maturité. Toutefois, les données à notre disposition ne nous permettent pas de déterminer dans quelle mesure le savoir que le candidat expose dans une composition influe sur l'évaluation de l'enseignant examinateur.

Comme l'indique le tableau V.19, l'examen des épreuves de Suisse alémanique fait ressortir, parmi les 19 catégories, les trois catégories 5.2 Créer-évaluer (D 30,9%), 4.3 Créer-reconnaître (D 17,8%) et 6.3 Créer-construire (D 19,5%).

Tableau V.19: Analyse des exigences cognitives (catégories principales et sous-catégories, pondérées⁴⁰; double désignation; épreuves de maturité de Suisse alémanique)

Exigences cognitives Suisse alémanique (avec pondération)		Points de pondération (D)	Fréquence (D)
Rappeler	1.1 Créer – identifier	2	0.4%
	1.2 Créer – rappeler	8	1.9%
Comprendre	2.1 Créer – interpréter	2	0.5%
	2.2 Créer – illustrer	30	6.6%
	2.3 Créer – classer	1	0.2%
	2.4 Créer – résumer	6	1.4%
	2.5 Créer – déduire	7	1.6%
	2.6 Créer – comparer	25	5.6%
	2.7 Créer – expliquer	14	3.2%
Appliquer	3.1 Créer – exécuter	4	0.8%
	3.2 Créer – appliquer	3	0.7%
Analyser	4.1 Créer – différencier	4	0.8%
	4.2 Créer – intégrer	16	3.6%
	4.3 Créer – reconnaître	79	17.8%
Evaluer / apprécier	5.1 Créer – vérifier	5	1.1%
	5.2 Créer – évaluer	138	30.9%
Créer	6.1 Créer – développer	12	2.8%
	6.2 Créer – planifier	3	0.6%
	6.3 Créer – construire	87	19.5%
Total des points de pondération ⁴¹		446	100.0%

Les résultats du Tessin sont comparables à ceux de la Suisse alémanique (cf. tableau V.20) puisque tant les travaux du domaine Créer-rappeler (catégorie 1) que du domaine Créer-appliquer (catégorie 3) ne comportent que très rarement des consignes explicites (données de <1> D 3,8%, I 6.1%), qu'un cinquième environ des sujets ressortissent de la catégorie Créer-comprendre (D 19,1%, I 19.2%) et que la grande majorité des énoncés d'examen peuvent être attribués aux exigences cognitives les plus élevées (catégories 4 à 6; D 77,2%, I 74,7%). Cependant, le groupe les plus fourni est celui des données du domaine Créer-analyser, qui représente presque un tiers des données recensées, suivi du domaine Créer-crée, avec un quart des sujets et un cinquième «seulement» relevant du domaine Créer-évaluer.

⁴⁰ Sur la procédure de pondération et la double dénomination: cf. les explications données plus haut à la section 3.1.3.

⁴¹ Les points de pondération sont fonction du nombre des énoncés de toutes les classes.

Tableau V.20: Analyse des exigences cognitives (examens de maturité au Tessin)

Exigences cognitives Tessin (avec pondération)	Points de pondération (I)	Fréquence (I)	Groupes
Créer - rappeler (1)	10	6,1%	
Créer - comprendre (2)	31	19,2%	
Créer - appliquer (3)	0	0,0%	25,3%
Créer - analyser (4)	48	29,8%	
Créer - évaluer (5)	31	19,1%	
Créer - créer (6)	42	25,8%	74,7%
Total des points de pondération	162	100,0%	100,0%

En Suisse romande, la prépondérance des tâches d'évaluation (en allemand *Beurteilungsaufgaben*) est plus marquée encore qu'en Suisse alémanique: deux tiers en effet des énoncés ont été attribués au domaine Créer-évaluer (niveau 5). Il ressort que la grande majorité des sujets – et dans une proportion notablement plus élevée (+14%) que dans les deux autres régions linguistiques – ressortissent de la moitié supérieure des niveaux d'exigences (niveaux 4 à 6, F 91,3%).

Tableau V.21: Analyse des exigences cognitives (examens de maturité en Suisse romande)

Catégories cognitives principales Suisse romande (avec pondération)	Fréquence (F)
Créer – rappeler (1)	0.2%
Créer – comprendre (2)	8.5%
Créer – appliquer (3)	0.0%
Créer – analyser (4)	24.2%
Créer – évaluer (5)	66.1%
Créer – créer (6)	1.0%
Total des catégories principales (pondérées ⁴²)	100.0%

3.1.4 Comparaison avec les tâches du module C

A la différence des examens de maturité, les tests en langue première du module C ont surtout ciblé la compréhension de l'écrit; il est donc possible de se prononcer sur les compétences dont témoignent les étudiants testés pour leur compréhension générale et détaillée de l'écrit, la compréhension du lexique, ainsi que dans les compétences de grammaire et d'orthographe (compétences sur le fonctionnement de la langue). S'agissant de la production écrite, le test C n'a proposé que quelques tâches isolées (telles que donner un intitulé à un paragraphe, reformuler des phrases) et qui n'étaient d'ailleurs que de brèves opérations. Les résultats des tests ne peuvent donc être mis en parallèle avec les formes de composition des épreuves de maturité.

Les épreuves écrites en langue première analysées dans le sous-projet D1, ont mis l'accent sur la production écrite. La capacité de lecture examinée dans le module C est cependant un prérequis indispensable pour qui veut parvenir à créer le produit complexe de la composition à la maturité. En effet, le travail de l'examen commence par une lecture et une compréhension correctes de l'énoncé et des textes soumis. C'est sur cette base seulement que peuvent se construire la réflexion individuelle et s'élaborer la production du texte personnel.

Il existe d'autres différences encore entre les deux situations d'exigences: dans le module C, les assertions sur la compréhension de l'écrit se réfèrent toujours à des textes documentaires de type scientifique. A la maturité, la compréhension de l'écrit attendue concerne des textes littéraires dans une proportion de 40 à 60%. Or, dans les textes littéraires, la production du sens s'effectue sur plusieurs niveaux. Ce processus complexe suppose que le lecteur saisisse les constellations d'images, les notations symboliques, les oppositions, les condensations (et d'autres phénomènes esthétiques) d'une manière qui n'a rien à voir avec le simple enregistrement

⁴² Sur la procédure de pondération et la double dénomination: cf. les explications données plus haut à la section 3.1.3

d'informations. La lecture interprétative de textes esthétiques et la lecture intelligente de textes scientifiques spécialisés font donc appel à des compétences partielles différentes dans le grand domaine de la compréhension de l'écrit. Les textes documentaires qui sont à la base d'à peu près 20% des épreuves de maturité ne font d'ailleurs pas partie du même genre de textes que les documents universitaires employés dans le module C: les premiers (module D) sont généralement des textes journalistiques ou des discours politiques, et la langue académique spécialisée n'y intervient que rarement (4%).

Dans la partie de réflexion sur la langue des tests C, on a examiné les prestations langagières en orthographe, ponctuation, grammaire et lexique: ces composants sont un matériau essentiel pour la production de textes écrits.

3.1.5 Comparaison avec les résultats du sous-projet A4⁴³

Comparer les résultats du sous-projet D1 (appréciation des compositions de maturité) avec ceux du sous-projet A4 (enquête auprès des enseignants) revient à se demander si les épreuves écrites de la maturité évaluent effectivement les performances, compétences et connaissances que les professeurs de nos hautes écoles attendent des nouveaux étudiants. En second lieu, il s'agit de voir si l'analyse des performances de fin de gymnase manifeste déjà des signes avant-coureurs des lacunes de connaissances et de compétences observées par les professeurs chez leurs étudiants des premiers semestres.

Les membres du corps professoral universitaire interrogés accordent en général la plus haute importance à la langue première comme discipline fondamentale du gymnase. Ils sont 60% en effet à attendre des gymnasiens des connaissances et des compétences solides en langue première, et 22% comptent sur la connaissance des bases essentielles (question 1). Ces données représentent les plus hautes valeurs de toutes les attentes formulées quant aux disciplines enseignées dans les gymnases. En observant le détail des formulations par lesquelles les enseignants font part de ce qu'ils considèrent comme des prérequis (de savoirs et de savoir-faire), on dresse le bilan suivant: avec 38%, la maîtrise de la langue est en tête des attentes, suivie de la maîtrise de la grammaire (10%), de l'orthographe (4,3%) et de la syntaxe (3,6%). La connaissance de la ponctuation n'est mentionnée que par 1% des professeurs interrogés. A peine 6% attendent des connaissances spécifiques de la discipline «langue première», telles que l'histoire de la littérature, 3% citent les méthodes d'analyse et d'herméneutique – précisions que ces mentions émanent en général de professeurs de germanistique ou d'autres langues.

La maîtrise de la langue à un niveau élevé est définie dans le détail par les formulations individuelles suivantes: «maîtrise de la langue du point de vue stylistique», «aisance dans la langue et dans l'écriture», «capacité d'expression écrite et orale de haut niveau», «aptitude à saisir d'une manière linguistiquement nuancée des contenus historiques», «expression écrite et orale en mesure de combiner le registre standard et la langue spécialisée», «sûreté stylistique». S'agissant plus précisément de la production de textes, les attentes suivantes figurent dans le sous-projet A4: «écriture d'une tenue scientifique», «bases de la composition», «rédaction sans faute de textes en langue première», «savoir rédiger des textes argumentatifs», «précision de la formulation», «vocabulaire étendu», «capacité de rédiger ses propres textes», «capacité de décrire le fonctionnement de la langue de façon correcte», «rédiger des textes structurés». La compréhension de l'écrit a donné lieu aux formulations suivantes: «exactitude de la lecture des textes», «connaissances sur l'élaboration d'une littérature spécialisée», «aptitude à saisir et à rendre des contenus complexes», «maîtrise de la langue même en face de textes complexes».

En mettant en parallèle ces attentes, formulées pour la discipline gymnasiale «langue première», et les épreuves écrites de maturité, on constate que par sa nature même, l'examen écrit de langue première va dans le sens de ce que souhaitent les enseignants des hautes écoles. Quand ils produisent un texte, la «composition», les

⁴³ Une évaluation comparative avec les résultats des sous-projets A1 et A3 ne pourra se faire que sur la base d'autres travaux plus approfondis.

bacheliers et bacheliers réfléchissent à un contenu – le thème imposé ou choisi – et manifestent aussi leurs capacités langagières générales dans les domaines «maîtrise de la langue», «expression écrite» et «aisance stylistique» comme leurs connaissances pratiques de grammaire, d'orthographe et de ponctuation en particulier.

Selon l'analyse des formes de composition exigées (cf. section 3.1.2.1), le genre du texte discursif représente en Suisse alémanique 24% de toutes les compositions (Erörterung, Argumentation), en Suisse romande 53% même (dissertation, texte argumentatif). Ces chiffres rejoignent les résultats de l'analyse des exigences cognitives (section 3.1.3): le niveau d'exigence du Créer-évaluer y représente un cinquième de l'ensemble des données d'examen au Tessin, en Suisse alémanique un tiers, et en Suisse romande deux tiers. D'après la sous-catégorie prépondérante (5.2), les bacheliers et les bacheliers doivent manifester qu'ils savent prendre position et apprécier une thèse en utilisant des arguments pour et contre. Ces résultats du sous-projet D1 attestent que les examens vérifient tout de même des formes d'écriture argumentative – prélude à l'écriture scientifique du moule académique – dans le sens de ce que souhaitent les universitaires. En effet, les attentes des professeurs en matière de compréhension de l'écrit sont satisfaites dans la mesure où presque tous les examens de maturité en langue première s'appuient sur un ou plusieurs textes de départ dont la compréhension exacte sert de base à la composition demandée.

Les attentes plus spécifiques des enseignants de langues trouvent aussi satisfaction sur un point particulier: la plupart des épreuves écrites soumettent des textes appartenant au domaine littéraire (D 41%, F 62%, I 57%) et portent ainsi sur l'application des connaissances attendues en méthodologie interprétative et herméneutique. Par contre, les examens ne vérifient guère – ou pas du tout – «la vue d'ensemble des époques les plus importantes de l'histoire de la littérature».

Interrogés sur les lacunes constatées dans le bagage des nouveaux étudiants, les professeurs font les constats suivants: en matière de maîtrise de la langue, le terme de pauvreté apparaît chez 22 enseignants – représentants tous d'autres domaines d'études que les langues. Il est intéressant de noter que cette lacune est observée par 11 des 16 professeurs de droit ayant participé à l'enquête. Une mauvaise maîtrise de la grammaire figure dans les réponses de 14 professeurs (essentiellement d'histoire et de germanistique), l'orthographe fautive est citée par 9 enseignants, et dans 4 cas, on regrette la syntaxe non maîtrisée et la ponctuation incorrecte. S'agissant de lettres, 5 germanistes et anglicistes regrettent l'absence de connaissances d'histoire de la littérature, 4 (des mêmes disciplines), les lacunes théoriques sur l'interprétation des textes littéraires, et deux professeurs de langues demandent de meilleures connaissances de linguistique. Les difficultés éprouvées dans la rédaction de textes scientifiques ainsi que l'utilisation de la littérature spécialisée ne sont mentionnées qu'une fois (pour chaque cas).

Partant des attentes que manifeste ce relevé des lacunes dans le bagage des nouveaux étudiants, on peut constater, ici également, une concordance avec les connaissances et compétences que les examens de maturité entendent évaluer, soit précisément, par la composition, la maîtrise de la langue (grammaire, syntaxe, orthographe et ponctuation incluses). Il y a lieu de relever que la note à l'épreuve écrite de maturité en langue première était insuffisante dans 18% des cas pour l'année 2007. Il est certain que l'on constatera de telles lacunes aussi longtemps que le système de calcul permettra la compensation des notes insuffisantes par des points acquis dans d'autres branches.

Outre les compétences et connaissances susmentionnées, les épreuves écrites de maturité en langue première examinent aussi implicitement les compétences transverses «penser de façon critique» et aussi, en partie, «développer de nouvelles idées» (questions 4 et 5 dans le sous-projet A4). La première de ces compétences entre en jeu dans les formes de composition qui font intervenir l'argumentation, la seconde dans les énoncés ressortissant des catégories cognitives 6.1 «développer» et 6.3 «construire». En Suisse alémanique, cette condition est remplie par 3% des sujets (pour la catégorie 6.1 «Créer-développer») ou même par 20% des sujets proposés (pour la catégorie 6.3 «Créer-construire»).

Les contenus d'examen recensés par D1 (auteurs, thèmes de société, questions éthiques, etc.) ne font pourtant pas partie de l'horizon des attentes formulées par les professeurs de toutes les disciplines ayant participé à

l'enquête. Ils correspondent, en revanche, au second objectif général de la formation gymnasiale suisse (RRM 95, art. 5), qui est de préparer les gymnasiennes et gymnasiens à assumer des responsabilités au sein de la société actuelle, lequel objectif, on le sait, n'était pas spécialement étudié par EVAMAR II.

En complément, on mentionnera ici les attentes spécifiques dont ont fait part les professeurs de germanistique dans le cadre du sous-projet A4 et on les mettra en parallèle avec les compétences évaluées dans l'épreuve écrite d'allemand à la maturité. Ce point concerne donc une aptitude plus spécifique (et non plus générale) aux études supérieures.

Interrogés sur les exigences posées aux futurs étudiants de germanistique, les enseignants de cette discipline ont indiqué explicitement leurs attentes de connaissances et de compétences (cf. chapitre II, section 3) dans les formulations suivantes:

- «premières connaissances de l'histoire littéraire» jusqu'à «vue d'ensemble des œuvres et des époques principales de l'histoire littéraire»;
- «expériences personnelles de l'analyse de textes» et «connaissance de base de quelques méthodes d'herméneutique» jusqu'à «compétence d'interprétation»;
- «maîtrise de notions fondamentales de linguistique et de science de la littérature»;
- «connaissance des types de textes, des formes de médias et de communication, conscience de la dimension diachronique de la langue»;
- «connaissance de la grammaire (espèces de mots, syntagmes, morphologie, syntaxe; déclinaison du substantif et de l'adjectif, distinction entre verbes forts et verbes faibles)»;
- «base de la structure de la langue allemande, terminologie grammaticale, capacité de reconnaître des régularités structurelles de la langue et de les formuler»;
- «connaissance des règles essentielles de la poésie occidentale, modèles importants de théorie de la lecture, connaissances et pratique».

A l'instar de leurs collègues d'autres disciplines scientifiques, ces professeurs ont également mentionné une bonne compréhension générale des textes, une bonne compétence rédactionnelle et une «parfaite connaissance de l'allemand standard» (cf. section 3.1.5).

Il est permis de constater que l'examen écrit de maturité d'allemand tel qu'il a été décrit dans la présente section *n'est guère en mesure* d'évaluer l'aptitude spécifique explicite à entreprendre des études de germanistique. En effet, parmi les quelque 40% d'énoncés d'examens qui se fondaient sur des textes en allemand des littératures de Suisse ou d'Allemagne, 26% consistaient en aphorismes (donc des extraits très brefs) et 62% étaient récents ou très récents (des 20^e ou 21^e siècles). Ces textes soumis aux élèves ne touchent donc que marginalement le vaste domaine scientifique de la germanistique. Il est aussi ressorti de l'analyse des examens écrits que ceux-ci ne posent pratiquement aucune exigence explicite en matière de connaissances sur les époques et qu'on ne demande aux candidats ni d'exposer des savoirs issus de la théorie des genres littéraires, de l'herméneutique ou de la grammaire, ni de connaître les généralités de l'histoire littéraire ou les bases des sciences de la littérature et de la linguistique. On en conclut que les examens écrits de maturité n'évaluent que dans une très faible mesure les connaissances qui devraient, du point de vue des professeurs concernés, avoir été acquises par les étudiants qui commencent en germanistique. Il faut cependant préciser qu'EVAMAR II n'a porté pas porté sur les examens oraux: on ne peut donc pas juger dans quelle mesure les épreuves orales de langue première touchent de plus près à ces connaissances et compétences qui sont des prérequis aux études de germanistique.

3.1.6 Appréciation globale des examens de maturité en langue première en Suisse

En Suisse, les examens écrits de maturité en langue première se distinguent par leur extraordinaire diversité: il semblerait que chaque canton, chaque école, chaque enseignante ou enseignant même, ait sa propre idée de la manière de composer cet examen et des conditions permettant de vérifier si les bacheliers et bacheliers possèdent la «maturité nécessaire aux études universitaires».

Le corps enseignant suisse est généralement d'accord sur deux points: la majeure partie des textes sont dans leur langue originale (il s'agit rarement de traductions) et les auteurs féminins sont rarement choisis. Les enseignantes et enseignants ont plutôt tendance à choisir un ou deux textes fondamentaux du 20^e ou 21^e siècles sur un sujet aussi récent que possible ou un thème intemporel. Ils choisissent en l'occurrence dans 40 à 60% des cas des textes littéraires et seulement dans 10 à 20% des cas des textes documentaires. Dans les autres cas, il s'agit de textes dits «philosophiques» (D allemand 12%, F français 14%, I italien 30%).

Le niveau d'exigence de trois quarts au moins des examens écrits de maturité se situe dans la partie supérieure (niveaux 4-6) des exigences cognitives (D 77%, F 91%, I 75%). Pour relever les principales différences régionales, on constate une nette prédominance du niveau d'exigence «création/critique» en Suisse romande (deux tiers), alors qu'elle ne vient qu'en troisième place au Tessin (un cinquième). En Suisse alémanique, la répartition est plus équilibrée: la réflexion critique forme un tiers des exigences, l'analyse un quart et la créativité un autre quart. Dans l'ensemble, le type d'examen de la composition de quatre heures relève clairement de la catégorie cognitive «Créer».

Sur le plan du contenu, les thèmes soumis aux élèves se réfèrent au monde actuel, plus particulièrement à celui des jeunes, et mettent l'accent sur la diversité des opinions et les débats de société ou les controverses sur des questions existentielles.

3.2 Mathématiques

L'analyse a porté sur 168 examens écrits de maturité: 101 de Suisse alémanique, 61 de Suisse romande et 6 de Suisse italienne. A peu d'exceptions près, les énoncés étaient rédigés dans la langue scolaire de la région. Dans le lot figuraient aussi certaines épreuves destinées aux examens d'option spécifique (OS) et d'option complémentaire (OC). La proportion des examens d'OS était plus élevée en Suisse romande.

3.2.1 Les instruments d'analyse

L'analyse du contenu des épreuves s'est fondée sur la grille de compétences élaborée dans le cadre du module B.

Pour l'analyse du potentiel d'activités cognitives des exercices, il a fallu adapter l'instrument décrit dans la section 1.2 aux mathématiques, comme l'indique le tableau V.22 ci-dessous.

Tableau V.22: Modèle des potentiels d'activités cognitives pour les mathématiques

Catégorie	Processus cognitifs	Autres formulations	Définitions
1	Rappeler faire appel à un savoir pertinent de sa mémoire à long terme		
1.1	reconnaître	identifier	trouver dans sa mémoire à long terme un savoir en relation avec les données reçues
1.2	faire appel	remémorer	faire appel à des connaissances pertinentes retenues par la mémoire à long terme
2	Comprendre saisir des contenus d'enseignement et en dégager du sens (oralement, par écrit, par une représentation graphique)		
2.1	interpréter	éclaircir, expliquer, paraphraser, reformuler, présenter autrement, traduire, transposer, retravailler	transposer des données reçues en modifiant la présentation
2.2	donner des exemples, illustrer, montrer de façon paradigmatique	clarifier, illustrer, représenter de façon imagée	trouver un exemple ou une situation imagée qui illustre une notion ou un principe
2.3	classifier	catégoriser, subsumer	décider de l'appartenance d'une instance à telle ou telle catégorie (de notions ou de principes)

Catégorie	Processus cognitifs	Autres formulations	Définitions
2.4	résumer	abstraire, généraliser	résumer un chapitre important ou les éléments principaux
2.5	déduire	conclure, inférer, extrapoler, interpoler, prévoir, estimer	déduire des conclusions logiques
2.6	comparer	contraster, opposer, ranger, rechercher des analogies	comparer diverses manières de résoudre un problème
2.7	expliquer	construire des modèles et des mécanismes	modéliser, utiliser des modèles mathématiques connus
3	Appliquer appliquer ou utiliser une procédure dans une situation donnée		
3.1	exécuter	–	appliquer une procédure à un problème déjà connu
3.2	implémenter	appliquer, faire usage de	appliquer une procédure à un problème inconnu
4	Analyser diviser une donnée reçue en ses éléments constitutifs, définir les rapports mutuels de ceux-ci et la relation qu'ils entretiennent avec la structure ou l'objectif d'ensemble		
4.1	différencier	distinguer, séparer, accentuer, sélectionner	discerner des éléments pertinents (ou importants) et d'autres qui ne le sont pas au sein des données reçues
4.2	organiser	trouver une logique interne, intégrer, esquisser les contours, analyser, structurer	déterminer comment des éléments prennent place et interagissent au sein d'une structure
4.3	se référer à quelque chose	déconstruire	résoudre un nouveau problème en le ramenant à un problème connu
5	Evaluer émettre un jugement sur la base de critères et de standards		
5.1	vérifier	mettre des éléments en relation, détecter des erreurs, vérifier, tester	vérifier ou contredire des affirmations
5.2	examiner de façon critique	apprécier	reconnaître si une procédure est adaptée à un problème donné
6	Créer réunir des éléments pour former un ensemble fonctionnel cohérent; articuler des éléments selon une nouvelle structure		
6.1	développer, élaborer	poser des hypothèses	modéliser, élaborer de nouveaux modèles mathématiques
6.2	planifier	donner forme	élaborer un mode de faire pour résoudre un problème posé
6.3	produire	construire	inventer un nouveau procédé ou processus

3.2.2 Les contenus des épreuves d'examen en mathématiques

3.2.2.1 Démarche

Les contenus des épreuves de maturité en mathématiques ont été classés sur la base de la grille de compétences développée dans le module B. Le programme de mathématiques est subdivisé en douze chapitres, dont la liste est donnée ci-dessous.

AF	Calcul analytique/fonctions
AI	Calcul analytique/calcul infinitésimal
AR	Calcul analytique/suites et séries
GT	Géométrie/trigonométrie
GG	Géométrie/géométrie élémentaire et analytique
GV	Géométrie/calcul vectoriel et géométrie vectorielle
SW	Stochastique/calcul des probabilités
SK	Stochastique/analyse combinatoire

SS	Stochastique/statistique
XM	Divers/théorie des ensembles
XL	Divers/logique
XA	Divers/algèbre élémentaire

Chaque exercice d'examen de maturité a reçu en principe un total de cinq points, répartis sur les chapitres prépondérants. Pour les exercices dont la résolution rapportait un nombre de points notablement plus élevé que pour d'autres, on a augmenté les points en conséquence. Un assez grand nombre d'épreuves de maturité n'indiquent aucun barème pour la résolution des problèmes; dans ces cas-là, on est parti du principe que le nombre des points possibles était identique pour chaque exercice.

La plupart des problèmes abordent, c'est une évidence, plusieurs domaines partiels des mathématiques; on s'est pourtant limité à attribuer des points aux seuls chapitres déterminants pour une question centrale. Il convient aussi de remarquer que les points n'ont été attribués que sur la base des énoncés, et non en fonction du temps nécessaire à l'accomplissement de la tâche. Pour des raisons de ressources, les didacticiens qui codaient les travaux n'ont pas été en mesure de résoudre eux-mêmes les problèmes posés en vue d'une ventilation précise des points selon les domaines. Ils ont donc attribué les points selon la procédure susmentionnée, puis les ont additionnés domaine par domaine et convertis en pourcentages, étant donné que les épreuves ne comportaient pas toutes le même nombre de questions. De cette manière, chaque épreuve de maturité a donné lieu à une répartition des fréquences (en pour-cent) pour chaque domaine de la grille de compétences (voir à ce sujet le rapport technique).

Pour résoudre les difficultés rencontrées lors de la catégorisation, on a adopté les solutions suivantes:

- Les très nombreuses études de fonctions ont valu 2.5 points au domaine AF et 2.5 points également au domaine AI. Maintes épreuves n'indiquaient pas le détail des tâches attendues des bacheliers et bacheliers, le canevas de l'étude de fonctions ayant vraisemblablement été convenu avant les examens. Dans certains cas, la donnée renvoyait à la démarche prévue dans un recueil de formules.
- Pour les problèmes d'optimisation de calcul différentiel, on a en principe attribué cinq points au crédit AI, sauf si d'autres chapitres des mathématiques se révélaient très importants pour la détermination de la fonction en discussion. Dans le cas de problèmes d'optimisation de calcul vectoriel (par ex. détermination du plus petit intervalle entre deux droites), les points ont été attribués au domaine partiel GV, bien que de tels problèmes puissent également être résolus par le biais du calcul différentiel.
- La résolution d'équations ou de systèmes d'équations n'a reçu des points que si l'épreuve de maturité comportait un exercice exclusivement réservé à la résolution d'une équation ou d'un système d'équations (dans ce cas au chapitre XA).
- Les énoncés qui comportaient des cercles, des sphères, des cônes ou des cylindres ont été rangés dans le domaine GG (géométrie élémentaire) et non GV (géométrie vectorielle).

3.2.2.2 Résultats

Le tableau V.23 présente les résultats de l'analyse. Les chiffres des pourcentages ont été arrondis à l'unité. La dernière colonne XX contient les domaines partiels des mathématiques qui n'entraient pas dans l'une des catégories définies plus haut (ex.: cryptologie, calcul modulaire, calcul faisant intervenir des nombres et fonctions complexes, algèbre linéaire, programmation linéaire, programmation informatique, physique, chimie, etc.).

Tableau V.23: Répartition (en pour-cent) des items selon les divers chapitres des mathématiques

	AF	AI	AR	GT	GG	GV	SW	SK	SS	XM	XL	XA	XX
Suisse alémanique	15%	28%	6%	3%	5%	16%	14%	6%	2%	0%	0%	2%	2%
Suisse romande	18%	25%	2%	1%	7%	7%	13%	3%	2%	0%	0%	2%	21%

Tessin	11%	19%	5%	0%	4%	10%	9%	0%	0%	0%	0%	0%	43%
Total	16%	26%	4%	2%	6%	13%	13%	5%	2%	0%	0%	2%	10%

Ce tableau fait apparaître qu'à l'échelle nationale, les énoncés les plus fréquents sont ceux de calcul infinitésimal (AI), suivis des études de fonctions (AF). La situation d'ensemble n'est pas la même dans les trois régions linguistiques: en Suisse alémanique, les problèmes de calcul vectoriel (GV) sont presque aussi fréquents que ceux de calcul des probabilités (SW), tandis que la Suisse romande réserve une place plus limitée aux données de calcul vectoriel. Dans les épreuves de Suisse romande, un nombre élevé d'item – au Tessin, même très élevé – ne pouvaient être intégrés à l'une des catégories de la grille (XX). Cette particularité s'explique par la proportion, plus grande en Suisse romande et au Tessin qu'en Suisse alémanique, d'épreuves écrites destinées à l'option spécifique (OS). Il existe aussi des champs et des contenus qui n'ont pas été évalués lors des examens: ainsi, les énoncés touchant à la théorie des ensembles (XM) ou à la logique (XL) n'apparaissent pas (ou: n'apparaissent plus) comme des énoncés autonomes à la maturité, ou alors très rarement. Des tâches centrées sur l'algèbre élémentaire (XA) ne représentent que 2% du total. Il s'agit principalement de problèmes demandant la résolution d'équations ou de systèmes d'équations. Les autres catégories peu représentées sont celles des items touchant aux suites et aux séries (AR), à la trigonométrie (GT) et à la statistique (SS).

Le tableau V.24 présente de façon synthétique les domaines partiels comme dans le module C, où ils ont servi de bases pour l'élaboration des problèmes.

Tableau V.24: Répartition (en pour-cent) des problèmes regroupés selon les principaux chapitres

	Analyse	Géométrie	Stochastique	Divers	XX
Suisse alémanique	49%	24%	22%	2%	2%
Suisse romande	45%	15%	18%	2%	21%
Tessin	35%	14%	9%	0%	43%
Total	46%	21%	20%	2%	10%

Il est manifeste que les problèmes d'analyse représentent presque la moitié des énoncés soumis aux candidats à la maturité en Suisse alémanique et en Suisse romande. Les problèmes de géométrie sont nettement moins nombreux en Suisse romande et au Tessin qu'en Suisse alémanique. L'écart provient surtout de la moins grande proportion de problèmes de géométrie vectorielle (GV). En revanche, la géométrie élémentaire et la géométrie analytique (GG) ont une importance légèrement supérieure en Suisse romande.

3.2.3 Les exigences cognitives

S'agissant de la détermination des exigences cognitives, il y a lieu d'appliquer aux mathématiques aussi les réserves décrites à la section 1.2, concernant la définition précise des niveaux d'exigence cognitive: c'est seulement le potentiel cognitif qui est susceptible d'être mis en œuvre. En outre, il n'était pas possible, pour des raisons d'organisation et de calendrier, d'obtenir les solutions des problèmes ni de les élaborer au sein de l'équipe d'EVAMAR II: il a fallu dès lors renoncer à affiner la catégorisation. On a donc effectué la détermination du niveau d'exigence cognitif sur la base des énoncés tels qu'ils figuraient dans les données d'examens que les écoles avaient mises à disposition.

Dans un premier temps, nous avons effectué un codage des consignes explicitement formulées («montrer»/ou «montrez», «calculer», «résoudre», «déterminer le terme de la fonction». Les questions du type «Quelle est la probabilité de... » ou «Quelle est l'équation de la sphère?» ont été reformulées avec des verbes à l'infinitif à valeur impérative: «Calculer la probabilité de...» ou «Déterminer l'équation de la sphère.». Dans un second temps, nous avons attribué les diverses tâches aux catégories cognitives adéquates.

Le tableau V.25 contient la liste complète des consignes explicites présentes dans les 162 épreuves écrites de maturité codées par EVAMAR II (Suisse alémanique et Suisse romande).

Tableau V.25: Exigences cognitives des items ou des problèmes

N°	Exercice ou problème/consigne	Catégorie d'après Anderson et al. (2001)	Taux en %
1	Etude de fonction (avec canevas explicite)	3.1	2
2	Etude de fonction (sans canevas explicite)	3.1	<0.5
3	Déterminer le domaine de définition d'une fonction	3.1	<0.5
4	Déterminer l'ensemble des valeurs	3.1	<0.1
5	Calculer les valeurs d'une fonction	3.1	1.5
6	Etablir le tableau des valeurs	3.1	<0.5
7	Poser l'équation de la fonction	3.1	3
8	Esquisser la représentation graphique	2.1	3
9	Etablir une correspondance entre des équations de fonctions et leur représentation graphique	2.1	<0.1
10	Déterminer les équations des asymptotes	3.1	<0.1
11	Etudier le comportement à l'infini (avec x tendant vers $+$ ou $-$ infini)	3.1	<1
12	Déterminer les zéros	3.1	<1
13	Déterminer les points d'intersection de deux représentations graphiques	3.1	<1
14	Calculer la dérivée d'une fonction	3.1	1
15	Calculer la pente/l'angle de pente	3.1	<0.5
16	Définir le tableau de variation	3.1	<0.1
17	Définir le tableau de courbure	3.1	<0.5
18	Rechercher les maxima/les minima	3.1	<0.5
19	Rechercher les points d'inflexion	3.1	<0.5
20	Déterminer l'équation d'une tangente à la courbe en un point donné	3.1	1.5
21	Déterminer l'équation de la normale en un point donné	3.1	<0.1
22	Calculer l'angle entre les représentations graphiques de deux courbes en un point	3.1	<0.5
23	Calculer l'angle entre des tangentes/entre des normales	3.1	<0.5
24	Déterminer la primitive d'une fonction	3.1	<0.5
25	Calculer une intégrale immédiate	3.1	<0.5
26	Calculer une intégrale plus complexe	3.1	<1
27	Calculer l'aire d'un domaine limité par une courbe	3.1	<4
28	Calculer l'aire d'un triangle	3.1	<0.5
29	Calculer la longueur d'une portion de courbe	3.1	<0.5
30	Déterminer le volume d'un solide de révolution	3.1	<2
31	Déterminer l'aire de la surface de révolution	3.1	<0.5
32	Résoudre un problème simple d'optimisation	4.3	<1
33	Résoudre un problème plus élaboré d'optimisation où une modélisation est indispensable	6.1	3
34	Résoudre des problèmes avec paramètres	3.1	4

N°	Exercice ou problème/consigne	Catégorie d'après Anderson et al. (2001)	Taux en %
35	Déterminer les coordonnées de points particuliers	3.1	<1
36	Démontrer un énoncé	5.1	<0.1
37	Appliquer une méthode numérique d'approximation	3.1	<0.5
38	Déterminer des lieux géométriques	3.1	<0.5
39	Problème de démonstration	5.1	<0.5
40	Calculer la vitesse et l'accélération d'un point	3.1	<0.1
41	Déterminer le polynôme de Taylor	3.1	<0.5
42	Calculer le temps de demi-vie ou le temps de doublement	3.1	<0.5
43	Calculer le taux de croissance ou le taux de décroissance	3.1	<0.5
44	Calculer les intérêts composés	3.1	<0.5
45	Poser une équation différentielle	3.1	<0.5
46	Résoudre une équation différentielle simple	3.1	<0.1
47	Résoudre une équation différentielle complexe	3.1	<0.5
48	Dessiner un diagramme des situations	3.1	<0.1
49	Déterminer les points d'équilibre	3.1	<0.5
50	Dessiner les trajectoires	3.1	<0.1
51	Etudier une forme quadratique	3.1	<0.5
52	Esquisser une conique	2.1	<0.5
53	Déterminer les coordonnées de points particuliers d'une conique	3.1	<0.5
54	Déterminer les coordonnées de points particuliers de fonctions à deux variables	3.1	<0.1
55	Déterminer les points selles	3.1	<0.1
56	Déterminer les extrema de fonctions à deux variables	3.1	<0.1
57	Déterminer les points d'intersection d'une courbe avec une surface	3.1	<0.1
58	Déterminer l'équation d'un plan tangent	3.1	<0.5
59	Calculer l'angle entre une courbe et son plan tangent	3.1	<0.1
60	Déterminer le centre de gravité d'une surface	3.1	<0.1
61	Calculer les termes d'une suite	3.1	<2.5
62	Donner une suite sous sa forme explicite	3.1	<0.5
63	Donner une suite par récurrence	3.1	<0.1
64	Calculer les sommes partielles	3.1	<1
65	Etudier la convergence d'une suite	5.1	<0.1
66	Déterminer le nombre des termes d'une suite remplissant une condition donnée	3.1	<0.5
67	Calculer la limite d'une suite	3.1	1
68	Calculer l'augmentation/la diminution annuelle en pour-cent	3.1	<0.5
69	Effectuer des résolutions dans un triangle rectangle	3.1	<0.5
70	Effectuer des résolutions dans un triangle quelconque	3.1	<0.1
71	Calculer l'aire d'un polygone	3.1	<0.5
72	Déterminer la représentation paramétrique d'une droite dans un espace à deux	3.1	<1.5

N°	Exercice ou problème/consigne	Catégorie d'après Anderson et al. (2001)	Taux en %
	dimensions ou dans un espace à trois dimensions		
73	Déterminer la représentation paramétrique d'un plan	3.1	<0.1
74	Déterminer l'équation explicite ou cartésienne d'un plan	3.1	2
75	Déterminer l'équation explicite ou cartésienne d'une droite dans un espace à deux dimensions	3.1	<0.5
76	Déterminer le vecteur directeur d'une droite	3.1	<0.1
77	Déterminer la position relative de deux droites dans l'espace	3.1	<0.5
78	Calculer les coordonnées du point d'intersection de deux droites	3.1	<0.5
79	Calculer les coordonnées du point d'intersection d'une droite et d'un plan	3.1	<2
80	Vérifier l'appartenance d'un point à une droite ou à un plan	5.1	<0.5
81	Calculer l'aire d'un triangle ou d'un parallélogramme	3.1	<1
82	Calculer le volume d'un tétraèdre, d'une pyramide ou d'un parallélépipède	3.1	<1
83	Calculer l'angle entre deux plans	3.1	<1
84	Calculer l'angle entre deux droites (ou entre deux vecteurs directeurs)	3.1	<1
85	Déterminer l'angle entre une droite et un plan	3.1	<1.5
86	Vérifier l'orthogonalité de deux droites	5.1	<0.5
87	Vérifier l'orthogonalité d'une droite et d'un plan	5.1	<0.5
88	Vérifier le parallélisme de deux droites	5.1	<0.1
89	Calculer les coordonnées d'un point	3.1	<1.5
90	Calculer la distance entre deux points	3.1	<1
91	Déterminer la distance entre deux droites	3.1	<0.5
92	Calculer la distance d'un point à une droite	3.1	0.5
93	Calculer la distance d'un point à un plan	3.1	<1.5
94	Calculer le milieu d'un segment	3.1	<0.5
95	Rechercher la valeur des paramètres dans l'équation d'une droite	3.1	<0.5
96	Déterminer la longueur d'un segment	3.1	<0.5
97	Déterminer l'équation d'un cercle	3.1	0.5
98	Calculer l'aire d'un cercle	3.1	<0.1
99	Déterminer l'équation (ou le centre/ou le rayon) d'une sphère	3.1	<2
100	Déterminer le plan tangent à une sphère	3.1	<1
101	Déterminer les coordonnées d'un point qui remplit certaines conditions	3.1	6
102	Calculer la distance entre deux points sur une sphère	3.1	<0.5
103	Dessiner un polyèdre	2.1	<0.1
104	Démontrer une propriété d'un quadrilatère	5.1	<0.5
105	Décrire une procédure	2.1	<0.5
106	Décrire un lieu géométrique	2.1	<0.1
107	Démontrer un énoncé géométrique	5.1	<0.1
108	Calculer la probabilité dans une expérience aléatoire à un niveau	3.1	2.5

N°	Exercice ou problème/consigne	Catégorie d'après Anderson et al. (2001)	Taux en %
109	Calculer la probabilité dans une expérience aléatoire à niveaux emboîtés	3.1	7
110	Calculer la probabilité conditionnelle	3.1	<1
111	Modéliser une situation à l'aide d'un schéma en arbre	2.1	<0.5
112	Calculer le nombre d'essais dans une expérience aléatoire à plusieurs niveaux	3.1	2
113	Démontrer que deux événements sont indépendants	5.1	<0.1
114	Résoudre un problème type d'analyse combinatoire	3.1	<2
115	Déterminer le nombre d'objets dans un problème d'analyse combinatoire	3.1	<0.5
116	Calculer la probabilité dans un problème type d'analyse combinatoire	3.1	3
117	Résoudre un problème élémentaire de statistique	3.1	<0.5
118	Déterminer la fonction de répartition	3.1	<0.5
119	Déterminer l'espérance mathématique	3.1	2
120	Déterminer la probabilité quand la distribution est donnée	3.1	<0.5
121	Calculer la variance	3.1	<0.1
122	Calculer l'écart-type	3.1	<0.5
123	Calculer un intervalle de probabilité	3.1	<0.1
124	Choisir et effectuer un test statistique	5.2	<0.5
125	Formuler l'hypothèse nulle	3.1	<0.5
126	Déterminer la taille de l'échantillon nécessaire	3.1	<0.1
127	Déterminer le domaine de rejet	3.1	<0.5
128	Calculer le risque d'erreur de premier ordre ou de deuxième ordre	3.1	<0.5
129	Vérifier une hypothèse	5.1	<0.5
130	Evaluer une stratégie de jeu	5.2	<0.1
131	Elaborer une stratégie de jeu	6.3	<0.1
132	Décrire une procédure	2.1	<0.1
133	Découvrir des variantes/découvrir un autre procédé	6.2	<0.1
134	Dénombrer les éléments d'un ensemble	3.1	<0.1
135	Expliquer verbalement des implications logiques	2.1	<0.1
136	Poser une équation	3.1	<1
137	Résoudre une équation	3.1	2
138	Résoudre une inégalité	3.1	<0.1
139	Poser un système d'équations	3.1	<0.1
140	Résoudre un système d'équations	3.1	<0.1
141	Résoudre une équation dans l'ensemble des nombres complexes	3.1	<0.1
142	Calculer la valeur en pour-cent	3.1	<0.5
143	Résoudre un problème standard en programmation linéaire	3.1	<0.1
144	Autres		<1.5

Il arrive souvent qu'une consigne en cache une autre, implicite; si on lit: «Déterminer l'aire entre les représentations graphiques de deux fonctions ...», le candidat doit la plupart du temps commencer par rechercher

les points d'intersection des deux représentations graphiques. Un autre énoncé peut donner comme tâche, donc comme consigne explicite, la détermination des points d'intersection des deux représentations graphiques. Pour la raison expliquée plus haut, il n'est pas possible de déterminer avec précision les parts quantitatives des diverses consignes qui apparaissent. Pour cette raison, nous nous limitons ici à relever un certain nombre d'observations:

Le plus couramment, les questions des données d'examens écrits sont du type: «Quelle est la probabilité que....?» Souvent, des tâches sont formulées comme ceci: «Calculer un point B tel que: ...» ou comme cela: «Pour quelles valeurs du paramètre a obtient-on:?».

Un exercice de maturité type comporte en moyenne quatre ou cinq consignes formulées explicitement. On observe que la plupart des tâches demandées correspondent au processus cognitif de l'application (catégories 3.1 et 3.2). Ces tâches d'application ont par exemple la forme: «Déterminer l'équation de la fonction ...» ou «Déterminer l'équation des coordonnées...» ou encore «Calculer la probabilité... ». La plupart du temps, il est malaisé de dire avec certitude si ces consignes se rapportent à la catégorie 3.1 (appliquer une procédure à un problème déjà connu) ou à 3.2 (appliquer une procédure à un problème inconnu; à ce propos, cf. aussi les remarques à la section 1.2). Les auteurs ont plutôt opté pour la catégorisation en 3.1.

Toute problème de maturité en mathématiques présuppose la présence de connaissances pertinentes. Donc, chaque donnée comporte aussi une part – plus ou moins importante – relevant de la catégorie Rappeler.

Très fréquemment aussi, des énoncés fournissant des équations de fonctions demandent au candidat d'en dessiner les représentations graphiques, ou d'autres invitent à représenter au moyen d'un diagramme en arbre une expérience aléatoire à niveaux emboîtés. Ces problèmes peuvent être attribués au domaine Comprendre, soit en 2.1, soit en 2.2.

Pour les études de fonctions, toujours très appréciées, on a différencié les cas où la donnée indiquait explicitement les tâches à effectuer («Déterminer le domaine de définition, trouver les zéros et les extrema ...») et ceux où les bacheliers et bacheliers devaient eux-mêmes se rappeler les étapes du canevas. Les problèmes avec études de fonctions ont aussi été rangés dans la catégorie 3.1.

Chaque examen de maturité en mathématiques ou presque comporte un problème d'optimisation: les plus simples ont été attribués à la catégorie 4.3. Les plus élaborés, qui supposent un travail préalable de transposition ou de modélisation, peuvent aussi être classés sous 6.1.

Les démonstrations telles qu'elles apparaissent surtout dans les examens de l'option spécifique «Physique et application des mathématiques (PAM)» ont été attribuées à la catégorie 5.1. On a classé de la même manière les consignes du genre «Vérifier que les deux droites sont gauches (ne sont pas situées dans un même plan)...» ou «Montrer que la droite g appartient au plan $E...$ » etc. S'agissant des consignes du type «Montrer que les droites se coupent dans l'espace ...», on a opté pour un classement dans la catégorie 5, bien qu'une tâche de ce genre puisse aussi être effectuée par le biais du calcul des points d'intersection (catégorie 3.1).

Les consignes du genre «Choisir un test statistique approprié, puis l'effectuer», qui apparaissent très rarement, ont été attribuées à la catégorie 5.2.

En relation avec le calcul des probabilités, il arrive de temps à autre qu'un exercice demande d'élaborer une stratégie de jeu. De tels énoncés ont été attribués à la catégorie 6.3.

Le tableau V.26 ci-dessous donne la récapitulation globale quantitative des exigences cognitives (potentielles) des données de maturité analysées.

Tableau V.26: Représentation quantitative des exigences cognitives potentielles

Catégorie	Taux
2.1 Comprendre: interpréter	< 5%
3.1 Appliquer: exécuter	85%
4.3 Analyser: rapporter à quelque chose	< 1%
5.1 et 5.2 Evaluer: vérifier et examiner de façon critique	< 5%
6.1 et 6.3 Créer: développer, élaborer, produire	< 5%

(<5% signifie que le taux est légèrement inférieur à 5%.)

3.2.4 Comparaison avec les items du module C

Il n'est guère possible de comparer les items des tests du module C avec les problèmes de maturité, car ces derniers requièrent un travail de résolution considérablement plus important que celui que les items ne peuvent et ne doivent demander. Par ailleurs, l'usage de la calculatrice de poche est généralement autorisé aux examens de maturité, mais non lors des tests. Malgré ces réserves, on peut relever les similitudes et les différences suivantes:

Autant dans les items des tests que dans les épreuves de maturité, on trouve des consignes telles que: «calculer les valeurs de la fonction...», «esquisser les représentations graphiques...», «établir une correspondance entre des équations de fonctions et leur représentation graphique...», «calculer les dérivées des fonctions...», «déterminer la primitive de...», «calculer la probabilité simple...». Malgré ce point commun, il n'est pas possible d'établir la corrélation entre le taux de réponses justes dans le test et la fréquence de l'exercice correspondant dans les consignes de maturité.

Les items de tests comme «esquisser les représentations graphiques ...» ont été mal résolus, manifestement parce que le terme de la fonction donnée n'était pas facile à comprendre. Les items concernant la primitive ont été résolus de façon moyenne seulement, quand bien même chaque examen de maturité (ou presque) comporte une recherche de la primitive, par exemple pour calculer l'aire d'un domaine limité par une courbe. On a enregistré des résultats satisfaisants pour d'autres questions du test, qui pourtant n'apparaissaient presque jamais comme tels dans les examens de maturité.

Il est possible de relever les observations suivantes: les données touchant à l'analyse représentent à peu près 50% des problèmes d'examens. Cependant, ce sont les items de ce domaine qui ont obtenus les moins bons résultats aux tests. La différence de taux de solutions justes dans les items de géométrie entre la Suisse alémanique et la Suisse romande pourrait s'expliquer par le fait que la Suisse romande comporte une moindre fréquence de problèmes de géométrie à l'examen. Des questions des tests touchant à des domaines non représentés aux examens de maturité (théorie des ensembles, logique) ont obtenu des résultats étonnamment bons, comme d'ailleurs des problèmes plutôt inhabituels liés à des figures géométriques ou à des fonctions à deux variables. Les données de maturité comportant de tels contenus sont très peu fréquentes et n'apparaissent, en fait, que dans les examens de l'option spécifique Physique et application des mathématiques (PAM).

3.2.5 Comparaison avec les résultats du sous-projet A4

Le tableau V.27 présente de façon synoptique les résultats de l'enquête des professeurs d'université selon le sous-projet A4 (cf. chapitre II, section 3). La première colonne est une liste récapitulative des réponses des professeurs à la question des prérequis les plus importants pour la fréquentation de leurs cours (question 2). Dans la deuxième colonne figurent les lacunes observées par les professeurs (question 3). Dans la troisième

colonne, nous avons introduit nos commentaires et les éléments de comparaison entre les affirmations des professeurs et les examens de maturité.

Tableau V.27: Comparaison des exigences des professeurs avec les examens de maturité

Prérequis, c'est-à-dire savoirs jugés importants pour les cours universitaires (parmi les chapitres de la discipline étudiés au gymnase (question 2))	Lacunes constatées chez les nouveaux étudiants parmi les connaissances et les compétences importantes (question 3)	Commentaire/comparaison avec les problèmes posés dans les examens de maturité
Généralités	Généralités	
être capable de formuler des énoncés mathématiques	esquisses d'états de faits mathématiques	trop vague
bases de la théorie scientifique		trop vague
conception de la démonstration		trop vague
	un savoir prêt à être appliqué de façon différenciée	trop vague
	connaissances générales des mathématiques	trop vague
	expression orale et écrite	EVAMAR II ne peut vérifier si la qualité de l'expression écrite est aussi prise en compte dans les épreuves de maturité (en mathématiques).
Logique	Logique	La logique n'apparaît pratiquement pas dans les épreuves de maturité.
connaissances de base en algèbre	Algèbre	
formules mathématiques	lecture des formules	Cet aspect est évalué surtout en relation avec les équations de fonctions ainsi que dans des énoncés concernant la description d'objets mathématiques (droite, plans) et comportant des vecteurs.
formule du binôme généralisé		Cet élément n'apparaît pas dans les épreuves écrites de maturité.
langage des formules par ex. les signes de somme		Chaque épreuves de maturité ou presque recourt au langage formalisé approprié.
savoir utiliser les variables		Pour résoudre la plupart des problèmes de maturité, il est nécessaire de savoir utiliser des variables. Un très grand nombre de problèmes des épreuves de maturité demandent de déterminer des paramètres de façon à satisfaire certaines conditions.
équations linéaires	équations linéaires et du deuxième degré	On trouve assez rarement des consignes demandant explicitement de résoudre des équations ou des systèmes d'équations. Les problèmes de calcul vectoriel et les problèmes de croissance ou décroissance supposent souvent des équations/des systèmes d'équations non explicités.
résoudre une ou deux équations à deux inconnues	équations à deux inconnues	Cf. Le commentaire ci-dessus sur les équations linéaires.
lois de puissances	lois de puissance	Cet élément n'apparaît pas dans les épreuves écrites de maturité.
Théorie des fonctions	Théorie des fonctions	

Prérequis, c'est-à-dire savoirs jugés importants pour les cours universitaires (parmi les chapitres de la discipline étudiés au gymnase (question 2))	Lacunes constatées chez les nouveaux étudiants parmi les connaissances et les compétences importantes (question 3)	Commentaire/comparaison avec les problèmes posés dans les examens de maturité
concept général de fonction, présentation de tableaux et de formules		Le maniement des fonctions représente une bonne moitié proposés à la maturité. Les examens de maturité comportent rarement des consignes explicites demandant d'établir des tableaux de valeurs. Pourtant, ceux-ci sont souvent indispensables à l'esquisse de représentations graphiques de fonctions, sauf si on se limite à des ponts particuliers (zéros, extrema, points d'inflexion)
Proportionnalité	Proportionnalité	Quand le principe de proportionnalité apparaît dans les examens de maturité, c'est le plus souvent en relation avec des calculs en pour-cent. Sinon, il s'agit surtout de décrire des phénomènes qui n'ont précisément pas un comportement proportionnel (processus de croissance et de décroissance exponentiels).
représentations graphiques de fonctions élémentaires		Plus du 90% des épreuves écrites contiennent un énoncé demandant de dessiner ou d'esquisser la représentation graphique d'une fonction.
étude de fonctions		Environ deux tiers des épreuves écrites comportent une donnée demandant une étude de fonction.
fonctions trigonométriques	fonctions trigonométriques (fonctions d'angle)	Ces fonctions sont très rares dans les énoncés de maturité actuels. Dans de nombreux cantons, ce chapitre ne fait plus partie du programme des deux dernières années.
fonctions exponentielles et logarithmiques	fonctions exponentielles et logarithmiques	L'analyse des examens de maturité n'a pas fait de distinction selon les types de fonction pour les problèmes demandant par exemple de déterminer l'équation de la fonction. Toutefois, les problèmes concernant les processus de croissance et de décroissance sont relativement rares.
minimum et maximum d'une fonction		Il n'existe pratiquement pas d'examen de maturité qui ne demande au moins une fois de rechercher le minimum ou le maximum d'une fonction.

Prérequis, c'est-à-dire savoirs jugés importants pour les cours universitaires (parmi les chapitres de la discipline étudiés au gymnase (question 2))	Lacunes constatées chez les nouveaux étudiants parmi les connaissances et les compétences importantes (question 3)	Commentaire/comparaison avec les problèmes posés dans les examens de maturité
Géométrie (analytique)	Géométrie (analytique) / Calcul vectoriel	
planimétrie (symétrie axiale, rotations)	planimétrie (longueur et mesure d'arcs)	La description analytique de figures géométriques est parfois demandée dans des questions d'examen de l'option spécifique PAM. Les problèmes de fonctions trigonométriques (qui comportent obligatoirement une mesure d'arc) sont plutôt rares.
Théorème de Thalès		N'apparaît pas explicitement dans les épreuves de maturité.
systèmes de coordonnées		Les systèmes de coordonnées sont la base permettant la représentation graphique des fonctions. On ne parvient pas à esquisser des représentations graphiques si on n'a pas compris ce qu'est un système de coordonnées.
Vecteurs	Vecteur (vecteur de connexion, longueur du vecteur)	Chaque épreuve de maturité ou presque comporte un problème de calcul vectoriel
analyse vectorielle		cf. ci-dessus
	géométrie analytique	trop vague
calcul intégral et différentiel	calcul intégral et différentiel	
calcul différentiel: dérivées	calcul différentiel: dérivées	Chaque épreuve de maturité ou presque comporte un problème de dérivée.
interprétations de la dérivée en physique		On en trouve seulement dans les problèmes de l'option spécifique PAM.
calcul intégral		Chaque épreuve de maturité ou presque comporte un exercice de calcul intégral.
théorie des ensembles		Elle n'apparaît pas dans les épreuves de maturité.
Stochastique	Stochastique	
calcul des probabilités	calcul des probabilités	Chaque épreuve de maturité ou presque comporte un exercice portant sur le calcul des probabilités.
bases de statistique	Statistique	Les problèmes de statistique sont plutôt rares.
différence entre rapport de corrélation et rapport de causalité		Cette distinction n'apparaît pas sous cette forme dans les épreuves de maturité.
suites et séries	Suites/Séries	
définitions de la suite géométrique et de la série géométrique	Définition de la suite géométrique/de la série	Un examen de maturité sur trois environ comporte un exercice traitant des suites et des séries. Il s'agit souvent de suites et séries géométriques, très fréquemment les termes de la suite sont présentés comme des longueurs de segments ou comme des aires de figures géométriques.
Remarques générales		
«En général, les connaissances d'algèbre linéaire sont limitées.»		L'algèbre linéaire ne fait pas partie du programme de la discipline fondamentale mathématiques.

Prérequis, c'est-à-dire savoirs jugés importants pour les cours universitaires (parmi les chapitres de la discipline étudiés au gymnase (question 2))	Lacunes constatées chez les nouveaux étudiants parmi les connaissances et les compétences importantes (question 3)	Commentaire/comparaison avec les problèmes posés dans les examens de maturité
«différences individuelles importantes»		trop vague
«On souhaiterait observer un meilleur sens des ordres de grandeur et moins de difficultés à conduire une pensée abstraite et analytique dans un contexte simple.»		trop vague On ne trouve pas de problèmes qui demandent d'indiquer seulement un ordre de grandeur du résultat.

Quelques-uns des domaines que les professeurs ont considérés comme lacunaires (algèbre linéaire, logique, etc.) ne font pas partie des chapitres indiqués par les gymnases comme matière évaluée à la maturité. Les chapitres de l'analyse, du calcul vectoriel, de la stochastique, sont très bien représentés dans les examens de maturité analysés, c'est-à-dire pratiquement dans toutes les données. Les problèmes de trigonométrie (tant les problèmes types que ceux qui portent sur les fonctions trigonométriques et leurs représentations graphiques) sont très rares, étant donné que ce domaine ne fait manifestement pas partie du programme des deux dernières années de gymnase. Or, c'est précisément sur ces deux dernières années que certains cantons, dont le canton de Berne, mettent l'accent principal aux examens de maturité.

3.2.6 Appréciation globale des examens de maturité en mathématiques en Suisse

Les examens de mathématiques présentent une grande diversité. En règle générale, les épreuves de maturité sont élaborées avec soin et beaucoup d'imagination.

Il existe cependant aussi un nombre considérable d'épreuves «standard» avec un problème par matière – étude de fonctions, calcul vectoriel, calcul de probabilité, calcul différentiel (problèmes d'optimisation) – et un exercice subdivisé en plusieurs items pour couvrir les autres domaines (suites et séries, algèbre élémentaire, etc.). Dans ces cas, la part de mémorisation est particulièrement importante sur le plan cognitif.

Il existe de grandes différences sur le plan des exigences cognitives (potentielles) tant entre les examens de mathématiques comme discipline fondamentale et les examens de l'option spécifique qu'entre les examens de type discipline fondamentale pratiqués par les diverses écoles. Pour préciser le propos sur ce point, il serait nécessaire d'examiner le détail des données d'examen et comparer les solutions des problèmes. Pour des raisons de calendrier, comme cela a déjà été mentionné, ce travail complémentaire n'a pu être effectué.

En ce qui concerne le potentiel d'exigences cognitives, la majeure partie des questions des épreuves de maturité appartiennent à la catégorie 3.1 («exécuter: appliquer une procédure à un problème déjà connu»). Le problème d'optimisation apparaissant dans presque chaque examen appartient généralement, à notre avis, soit à la catégorie 4.3 («résoudre un nouveau problème en le ramenant à un problème connu») soit à la catégorie principale 6 (Créer), lorsqu'il s'agit de problèmes d'optimisation plus élaborés nécessitant un important travail de modélisation.

3.3 Biologie

Pour la biologie, l'analyse a porté sur les épreuves de 65 examens de maturité de l'année 2007. Ce nombre restreint s'explique du fait que la biologie, en tant que partie de la discipline fondamentale du domaine des sciences expérimentales, ne fait pas l'objet d'un examen écrit dans la plupart des écoles. La plupart des examens analysés proviennent de l'examen de maturité pour l'option spécifique ou complémentaire. Les instruments d'analyse sont décrits ci-après.

3.3.1 Les instruments d'analyse

Toutes les questions des examens de maturité ont été classées selon leur contenu en six sous-domaines thématiques conformes à la grille de compétences en biologie (cf. chapitre III, section 3). Le tableau V.28 présente une vue d'ensemble de ces six sous-domaines.

Tableau V.28: Domaines thématiques pour l'analyse des épreuves de maturité en biologie

Domaine thématique (par analogie à des ouvrages standard en allemand pour l'enseignement de la biologie au degré secondaire II)					
Cytologie/Anatomie/Métabolisme	Traitement de l'information/Comportement	Immunobiologie	Génétique/Biologie du développement	Ecologie	Systématique/Evolution

La classification volontairement grossière en six domaines seulement s'est avérée payante lors du codage: la plupart des épreuves ont aisément pu être classées dans une catégorie. Souvent, il a été possible de s'inspirer des intertitres que les auteurs des examens avaient eux-mêmes placés en tête des groupes de questions.

L'instrument utilisé pour l'analyse du potentiel d'activités cognitives correspond à celui décrit à la section 1.2 (voir tableau V.1 ci-dessus). Ce modèle est largement connu et bien accepté en sciences naturelles. La spécification détaillée des catégories sur laquelle se sont fondés les travaux de codage est publiée dans le rapport technique, avec de nombreux exemples de questions et de problèmes concrets.

Il a fallu structurer encore un peu plus les examens pour l'analyse. Dans le cas le plus simple, le nombre maximal de points disponibles pour chaque question et sous-question des examens de biologie a été indiqué directement par l'auteur de l'examen, ce qui a été possible pour 58 sur 60 examens de Suisse alémanique. En Suisse romande, par contre, ces indications n'étaient disponibles que pour 4 examens sur les 10 analysés. Lorsque le nombre maximal de points était indiqué pour une question sans indication de la répartition des points par sous-question, on a divisé le nombre total de points par le nombre de sous-questions. Ce principe a été appliqué de façon automatique chaque fois que le barème effectif de l'examineur n'était pas connu.

Les épreuves de biologie ont été codées exclusivement sur la base des examens originaux écrits. On a volontairement ignoré les choix terminologies et l'énoncé spécifique des questions (surtout les verbes) comme critère prétendument sûr et objectif d'appartenance à une catégorie. La catégorisation se fonde bien davantage sur l'estimation des processus cognitifs nécessaires ou probablement activés pour répondre aux questions. Généralement, un simple examen de plausibilité permet d'estimer relativement précisément quels seront les processus en jeu. Un énoncé tel que «Expliquez de quelle manière l'ADN se réplique à l'aide d'enzymes» ne peut évidemment pas exiger d'un bachelier passant l'examen de biologie qu'il développe de façon plausible, sur la base de ses connaissances de biochimie, d'enzymologie et de la structure de l'ADN acquises au gymnase, le mécanisme de la réplication (une performance récompensée par le Prix Nobel en 1959). Selon toute vraisemblance, le verbe «expliquez» est employé au sens de «décrivez» comme un simple appel à la mémoire à long terme des candidats (catégorie 1.2). Dans d'autres cas, il peut être difficile de trancher si un fait, une structure, un processus ou une expérience mentionnés dans l'énoncé d'une question ont déjà été présentés sous une forme analogue aux élèves pendant les cours ou s'il s'agit plus probablement d'une nouvelle matière spécialement préparée par l'auteur de l'épreuve pour les besoins de l'examen. Pour la présente catégorisation, nous avons adopté pour critère de reconnaissance principal d'une nouvelle matière une forme de présentation cohérente et complète analogue à celle qu'on pourrait trouver dans un manuel, c'est-à-dire avec un degré élevé de cohésion textuelle. Par opposition, nous avons interprété un énoncé bref, construit autour de quelques mots clés isolés et dénués d'explications, comme un indice que les bacheliers étaient supposés reconstruire le reste du contexte sur la base des éléments vus au cours de leur formation.

De façon générale, nous avons codé comme tâches requérant un degré élevé de transfert les problèmes et questions ayant les principales caractéristiques suivantes:

- la question est tendanciellement placée à la fin d'un groupe de sous-questions, typiquement après plusieurs questions de la catégorie 1.2 (Appel à la mémoire à long terme);
- la question est énoncée au conditionnel: «Quel pourrait être le résultat?», «Pourquoi serait-il possible que ...?».

3.3.2 Les contenus des épreuves d'examen en biologie

Il s'agit d'examiner dans quelle mesure la répartition thématique des épreuves d'examens de maturité passés en Suisse en 2007 diverge d'une répartition «neutre» communément admise. Une telle répartition communément admise peut être déduite de la structure de trois ouvrages standard en allemand pour l'enseignement de la biologie au gymnase (degré secondaire II), plus précisément en se fondant sur le nombre de pages consacrées à chaque domaine thématique (voir tableau III.1). Pour estimer le volume relatif de chacun des six domaines thématiques de la grille de compétences (voir chapitre III, section 4.3) dans chacun des trois manuels standard de biologie, nous avons compté les pages correspondantes pour ensuite diviser le nombre ainsi obtenu par le nombre de pages total de l'ouvrage en question.

Il apparaît que le poids relatif de ces chapitres dans chacun des trois ouvrages standard de biologie étudiés est analogue (voir tableaux complets dans le rapport technique). Cette analogie confirme la plausibilité de l'approche consistant à utiliser ces pourcentages comme référence au sens d'une pondération communément admise.

La part moyenne des points attribués à un domaine thématique t par rapport au nombre total de points d'un examen de maturité de l'année 2007 a été estimée au moyen du calcul suivant:

- examen de maturité j (donné)
- question partielle d'examen Q_{ij} = question d'examen i dans le cadre de l'examen de maturité j (donnée)
- n_G = nombre total d'examens de maturité analysés = 65 (donné)
- z_j = nombre maximal de points disponibles pour l'ensemble de l'examen de maturité j (donné)
- a_{ij} = nombre maximal de points disponibles pour la question d'examen Q_{ij} (donné)
- M_t = ensemble de tous les a_{ij} pour les questions d'examen Q_{ij} appartenant au domaine thématique t (selon l'analyse de contenu; donné)
- part moyenne estimée du domaine thématique t par rapport au nombre total de points disponibles dans un

$$\text{examen de maturité 2007} = w_{tG} = \frac{1}{n_G} \sum_{j=1}^{n_G} \sum_{a_{ij} \in M_t} \frac{a_{ij}}{z_j}$$

Les résultats de l'analyse sont présentés dans le tableau V.29.

Tableau V.29: Poids moyen du domaine thématique t dans la note ou le nombre total de points disponibles

	Cytologie/ Anatomie/ Métabolisme	Traitement de l'information et comportement, regroupés avec l'immunobiologie	Génétique/ Biologie du développe- ment	Ecologie	Systématique/ Evolution	Divers	Total
Importance relative (en termes de pagination) des chapitres consacrés à ce domaine thématique dans trois ouvrages standard de biologie en langue allemande au niveau gymnasial (degré SII)	23.1%	24.2%	19.6%	17.3%	15.8%	0%	100.0%

	Cytologie/ Anatomie/ Métabolisme	Traitement de l'information et comportement, regroupés avec l'immunobiologie	Génétique/ Biologie du développe- ment	Ecologie	Systématique/ Evolution	Divers	Total
I Part moyenne estimée de ces domaines thématique par rapport	28.5%	21.9%	26.2%	10.2%	6.4%	6.9%	100.1%
II Pourcentages corrigés à partir de I après répartition proportionnelle de la catégorie «divers»	30.6%	23.5%	28.1%	10.9%	6.9%	0%	100.0%
Part des domaines thématiques dans les examens de maturité (pourcentages corrigés) par rapport à la répartition dans les ouvrages standard de biologie (définie ici comme valeur de référence à 100%)	132%	97%	143%	63%	44%		

En comparaison avec la pondération dans les manuels d'enseignement, un candidat ou une candidate à un examen de maturité moyen en biologie peut influencer son nombre total de points de manière particulièrement efficace avec des compétences dans les domaines Cytologie/Anatomie/Métabolisme et Génétique/Biologie du développement. Par contre, les domaines thématiques Ecologie et Systématique/Evolution sont nettement sous-pondérés dans les examens de maturité en biologie par rapport à la place qu'ils occupent dans les ouvrages de référence. Cela pourrait être lié au fait que la théorie de l'évolution passe souvent au second plan dans l'enseignement gymnasial. La négligence de cette thématique pourrait expliquer en partie que les items relatifs à la théorie de l'évolution ont été relativement mal résolus y compris par les bons bacheliers et les bonnes bachelières, si bien que nombre d'items de ce domaine thématique n'ont pas atteint la valeur discriminatoire minimale nécessaire pour entrer dans la sélection des items de biologie pour EVAMAR II (choix final des items pour l'étude principale; cf. chapitre IV, section 2.4).

3.3.3 Exigences cognitives

Les bacheliers et les bachelières sont-ils préparés aux exigences cognitives des problèmes d'examen à l'université? Tandis qu'il est difficile de répondre à cette question pour les disciplines également étudiées dans EVAMAR II des mathématiques et de la langue première – lesquelles préparent les élèves aux exigences de nombreuses disciplines d'études et pour lesquelles il faudrait analyser les examens en question en filtrant les éléments purement liés aux mathématiques et à la langue –, la comparaison est plus aisée en biologie. Cette discipline de maturité est plus que les deux autres conçue comme une préparation directe aux études de biologie à l'université et peut donc être directement évaluée à la lumière des examens universitaires en biologie. Il est possible de répondre à la question d'une préparation adéquate en comparant le potentiel d'activation cognitive de questions d'examen typiques des deux niveaux de formation. Une hypothèse souvent émise voudrait que les questions de biologie aux examens suisses de maturité 2007 présentent un potentiel d'activation cognitive plus élevé que les questions des examens universitaires. Cette hypothèse est généralement étayée par le raisonnement suivant: en raison du nombre élevé d'étudiants par cours et de la charge de travail importante qui en résulte pour le personnel enseignant, une pression structurelle s'exerce pour limiter le travail de correction, qui conduit les auteurs des examens à formuler une grande partie des questions dans le format – pratique à corriger – des questions à choix multiple. Alors que les questions ouvertes présentent elles un potentiel d'activation élevé en raison des capacités métacognitives et de transfert élevées qu'elles mobilisent, elles sont

pénalisées par la difficulté qu'il existe de définir un clé d'évaluation objective pour les corriger. Cette difficulté a d'une part l'inconvénient de rendre la correction plus attaquable, ce qui représente un certain risque pour les examinateurs au vu de l'énorme importance des résultats de l'examen pour chaque candidat. D'autre part, une clé de correction moins objective peut difficilement être partagée avec d'autres personnes. La personne chargée du cours n'aura donc guère de possibilité de déléguer une partie de son travail de correction si celui-ci s'appuie sur une clé de correction relativement peu objective.

Pour l'analyse, nous avons tout d'abord déterminé l'importance moyenne, en termes de note, des questions présentant un certain potentiel d'activation cognitive dans un examen de maturité. Le poids relatif de chaque type de questions a été estimé sur la base du calcul suivant:

- examen de maturité j (donné)
- question partielle d'examen Q_{ij} = question d'examen i dans le cadre de l'examen de maturité j (donnée)
- n_G = nombre total d'examens de maturité analysés = 65 (donné)
- z_j = nombre maximal de points disponibles pour l'ensemble de l'examen de maturité j (donné)
- a_{ij} = nombre maximal de points disponibles pour la question d'examen Q_{ij} (donné)
- M_i = ensemble de tous les a_{ij} pour les questions d'examen Q_{ij} présentant le potentiel d'activation p (selon l'analyse de contenu)
- poids moyen estimé des questions présentant un potentiel d'activation cognitive p à un examen de maturité =

$$w_{pG} = \frac{1}{n_G} \sum_{j=1}^{n_G} \sum_{a_{ij} \in M_p} \frac{a_{ij}}{z_j}$$

Les résultats relatifs aux valeurs w_{pG} sont présentés dans le tableau V.30.

Tableau V.30: Part du nombre total de points pouvant être obtenus par la résolution de questions d'une certaine catégorie dans un examen de maturité moyen en biologie.

Catégorie principale et sous-catégories	Définitions adaptées pour la discipline de la biologie dans EVAMAR II	Part du nombre total de points pouvant être obtenus par la résolution des questions de cette catégorie
	[d'après Anderson et al. (2001), p. 63–92]	w_{pG} en pour-cent
1	rappeler faire appel à un savoir pertinent de sa mémoire à long terme	
1.1	trouver dans sa mémoire à long terme un savoir en relation avec les données reçues	1.6%
1.2	faire appel à des connaissances pertinentes retenues par la mémoire à long terme	47.0%
2	comprendre saisir des contenus d'enseignement et en dégager du sens (oralement, par écrit, par une représentation graphique)	
2.1	transposer des données reçues en modifiant la forme de présentation	1.4%
2.2	trouver un exemple ou une situation imagée qui illustre une notion ou un principe	0.6%
2.3	décider de l'appartenance d'une instance à telle ou telle catégorie (de notions ou de principes)	1.3%
2.4	à partir de données reçues, trouver un sujet général ou résumer les éléments principaux (avec un processus personnel d'intégration) étudier les données reçues pour en tirer la quintessence, la pointe, «l'enjeu» et résumer (avec un processus personnel d'intégration)	0.5%

Catégorie principale et sous-catégories	Définitions adaptées pour la discipline de la biologie dans EVAMAR II	Part du nombre total de points pouvant être obtenus par la résolution des questions de cette catégorie
	[d'après Anderson et al. (2001), p. 63–92]	w_{pG} en pour-cent
2.6	reconnaître les convergences et les différences entre des idées, des objets et autres ordonner et classer des idées, des objets et autres au sein d'un même niveau hiérarchique (anglais: <i>mapping</i>)	5.9%
2.7	appliquer à des éléments que l'on découvre un modèle, appris auparavant, décrivant un rapport de cause à effet appliquer signifie: si le donné = le ou les effet(s): utiliser le modèle pour remonter aux causes si le donné = la ou les cause(s): utiliser le modèle pour conclure aux effets	20.6%
3	appliquer appliquer ou utiliser une procédure dans une situation donnée	
3.2	appliquer une procédure (étapes qui s'enchaînent ou qui ont pu être apprises) à un problème <u>non</u> connu	11.1%
4	analyser diviser une donnée reçue en ses éléments constitutifs, définir les rapports mutuels de ceux-ci et la relation qu'ils entretiennent avec la structure ou l'objectif d'ensemble	
4.1	discerner des éléments pertinents (ou importants) et d'autres qui ne le sont pas au sein des données reçues (sans processus personnel d'intégration)	0.9%
4.2	déterminer comment des éléments prennent place et interagissent au sein d'une structure structurer des données reçues et les transposer sous une autre forme (attribuer des sous-titres, construire un tableau, élaborer un diagramme en bâtons, un histogramme, un topogramme, etc.)	1.5%
4.3	à partir des données reçues, reconnaître le point de vue d'un auteur, les distorsions de perception et de jugement, sa représentation des valeurs ou ses intentions	0.2%
5	évaluer émettre un jugement sur la base de critères et de standards	
5.1	détecter des illogismes ou des erreurs au sein d'une démarche ou d'un produit; décider si la démarche ou le produit est en soi logique; déterminer l'effectivité d'une procédure (telle qu'elle est appliquée)	2.9%
5.2	déceler, en fonction de critères externes, les points faibles d'un processus ou d'un produit; définir la cohérence externe d'un produit; reconnaître si une procédure est appropriée à un problème donné	3.1%
6	créer réunir des éléments pour former un ensemble fonctionnel cohérent; articuler des éléments selon une nouvelle structure	
6.1	élaborer des hypothèses	1.3%
6.2	élaborer un mode de faire pour résoudre un problème posé	0.2%
6.3	inventer un nouveau produit	0.0%

Trois caractéristiques des examens de maturité 2007 en biologie se dégagent:

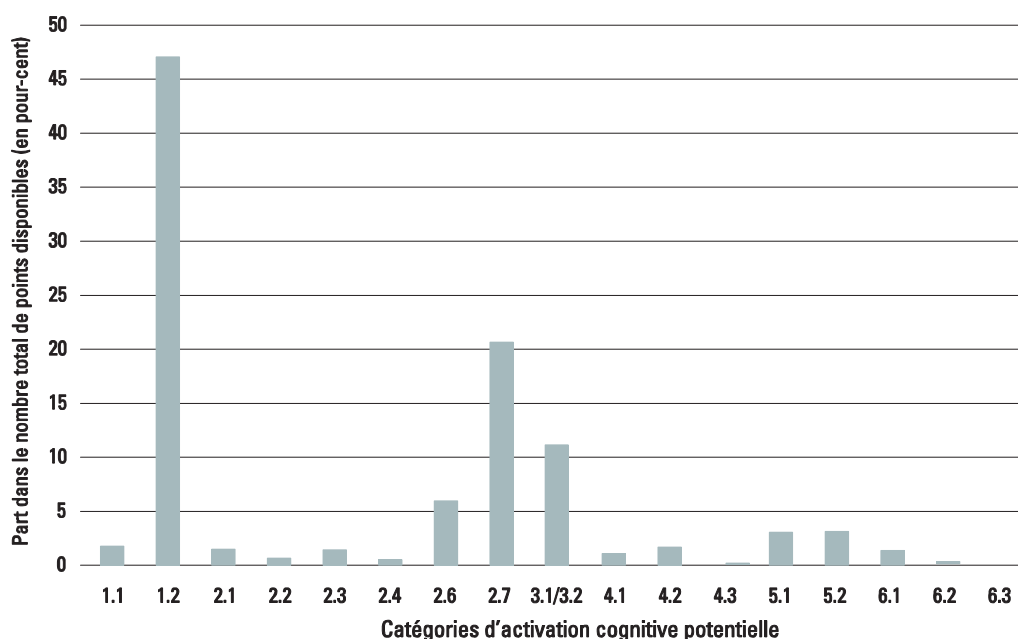
- a) Environ la moitié du nombre total de points (48,6%) d'un examen de maturité moyen en biologie est allouée pour des questions de connaissances pures (catégories 1.1 et 1.2) qui ne requièrent aucun processus cognitif au-delà du rappel d'informations enregistrées dans la mémoire.
- b) Les bacheliers et les bachelières qui ont une bonne maîtrise des quatre principaux processus cognitifs ci-après obtiennent déjà 84,4% du nombre total de points:
- faire appel à des connaissances pertinentes retenues par la mémoire à long terme (1.2),
 - ordonner et classer des idées, des objets et autres au sein d'un même niveau hiérarchique (anglais: *mapping*)(2.6),
 - appliquer à des éléments que l'on découvre un modèle, appris auparavant, décrivant un rapport de cause à effet (2.7),
 - appliquer une procédure (étapes qui s'enchaînent ou qui ont pu être apprises) à un problème non connu (3.1/3.2).

La maîtrise des 17 autres processus cognitifs n'apporte pour leur ensemble que 15,6% du nombre total de points. La répartition est donc très inégale.

- c) En vue d'estimer la contribution en points de questions qui activent potentiellement des processus de transfert et de métacognition (Georghiades, 2000), on peut additionner la dotation moyenne en points des catégories principales 3 à 6 (Anderson et al., 2001, p. 63). Les tâches rangées dans la sous-catégorie 2.7 activent elles aussi clairement des capacités de transfert au sens de la définition choisie dans le cadre d'EVAMAR II. Leur dotation moyenne en points doit donc être additionnée à la contribution en points recherchée. A noter que le processus cognitif de la comparaison (sous-catégorie 2.6) n'est pas compté ici comme un processus de transfert. Selon cette répartition, la part du nombre total de points pouvant être obtenue par la résolution de tâches exigeantes en termes de transfert et de métacognition est de 42%.

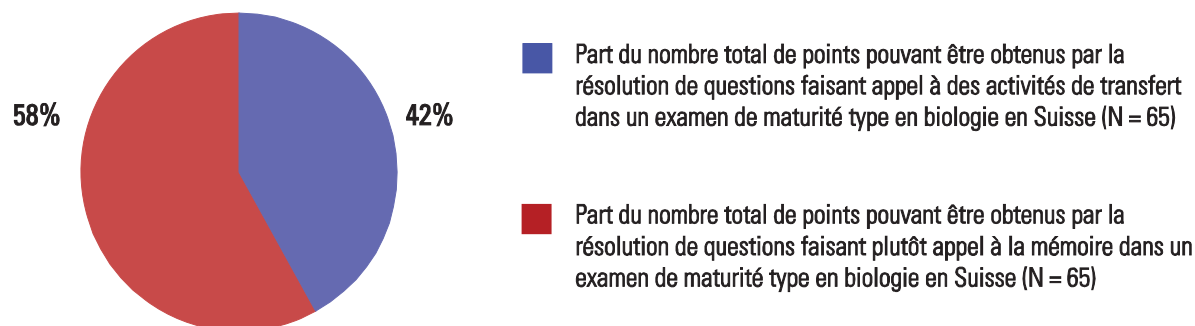
La figure V.1 montre que les parts du nombre total de points sont très souvent réparties de manière groupée. En répondant correctement aux questions des quatre catégories de tâches les plus importantes, il est déjà possible d'obtenir 84,4% du nombre total de points d'un examen de maturité moyen en biologie.

Figure V.1: Parts des différentes catégories dans le nombre total de points, N = 65
(54 examens de maturité, 9 examens de préparation à la maturité, 2 examens de biologie en tant qu'option complémentaire; D 58, F 4, I 3)



La figure V.2 montre par ailleurs qu'il n'est possible d'obtenir que 58% du nombre total de points en résolvant correctement toutes les tâches plutôt fondées sur la mémoire d'un examen de maturité moyen en biologie. Le reste des points doit être obtenu en faisant appel à des activités de transfert.

Figure V.2: Parts respectives des tâches de mémoire et de transfert (biologie, gymnase)



Pour examiner les examens universitaires en biologie, nous avons appliqué fondamentalement la même méthode de calcul que pour les examens de maturité. Le nombre total des examens analysés se limite toutefois à $n = 26$, car tous les examens ne nous ont pas été mis à disposition. Nous disposions de 19 examens du premier semestre des études de biologie à l'EPFZ et de 7 examens des deux premiers semestres à l'Université de Zurich. Comme l'échantillon de l'Université de Zurich était trop petit et que les échantillons des deux hautes écoles n'étaient pas randomisés, on a renoncé à comparer les deux écoles pour les considérer ensemble, chacun des deux établissements ayant la même pondération de 1.0 dans l'analyse globale (voir colonne de droite dans le tableau V.31). A l'EPFZ, les cours n'ont pas tous le même poids pour les notes. Il en a été tenu compte dans l'analyse. Comme nous n'avions cependant pas tous les examens des cours du 1^{er} semestre à notre disposition, nous avons réparti le poids total de l'EPFZ sur les trois cours analysés, proportionnellement à la pondération des notes ressortant du règlement d'examen pour les études de bachelor 2005/06. Les cours dont les types de questions ont pu être analysés sur la base d'un nombre relativement important d'examens n'ont bien entendu pas reçu un poids plus important dans l'appréciation finale que les cours pour lesquels un nombre d'examens plus limité était à disposition. Cela se traduit par la division par trois du poids des notes attribuées pour le cours «Bases de la biologie: biologie générale IA» à l'EPFZ. A l'Université de Zurich, chaque examen de module doit être passé avec une note suffisante. Les examens de modules ont donc tous le même poids. Etant donné que des examens de cinq modules étaient à disposition pour l'analyse, le poids total de l'Université de Zurich a été divisé par cinq, chaque module recevant une pondération de 0,2. L'estimation se fonde en détail sur le calcul suivant:

- examen universitaire j (donné)
- question partielle d'examen Q_{ij} = question d'examen i dans le cadre de l'examen universitaire j (donnée)
- n_H = nombre total d'examens universitaires analysés = 26 (donné)
- z_j = nombre maximal de points disponibles pour l'ensemble de l'examen universitaire j (donné)
- a_{ij} = nombre maximal de points disponibles pour la question d'examen Q_{ij} (donné)
- M_p = ensemble de tous les a_{ij} pour les questions d'examen Q_{ij} présentant le potentiel d'activation p (selon l'analyse de contenu)
- e_{ij} = nombre d'examens analysés du cours auquel se rapporte la question Q_{ij}
- d_{ij} = nombre d'enseignants dispensant le cours auquel se rapporte la question Q_{ij}
- w_{reg} = poids relatif du cours selon le règlement des examens, en admettant que le poids total des examens analysés est de 1.0 par établissement

- poids moyen estimé, en termes de note, des questions présentant un potentiel d'activation p à un examen

$$\text{universitaire} = w_{pH} = \sum_{j=1}^{n_H} \sum_{a_{ij} \in M_p} \frac{a_{ij}}{z_j} \cdot \frac{1}{e_{ij}} \cdot \frac{1}{d_{ij}} \cdot w_{reg}$$

On trouve une vue d'ensemble des examens universitaires analysés dans le rapport technique. Les résultats de l'analyse sont résumés dans le tableau V.31. Il montre la part du nombre total de points qui peut être obtenue par la résolution de tâches d'une certaine catégorie dans le cadre d'un examen moyen des deux premiers semestres à l'université.

Tableau V.31: Répartition du nombre total de points dans les examens universitaires en biologie par catégories de potentiel d'activation cognitive

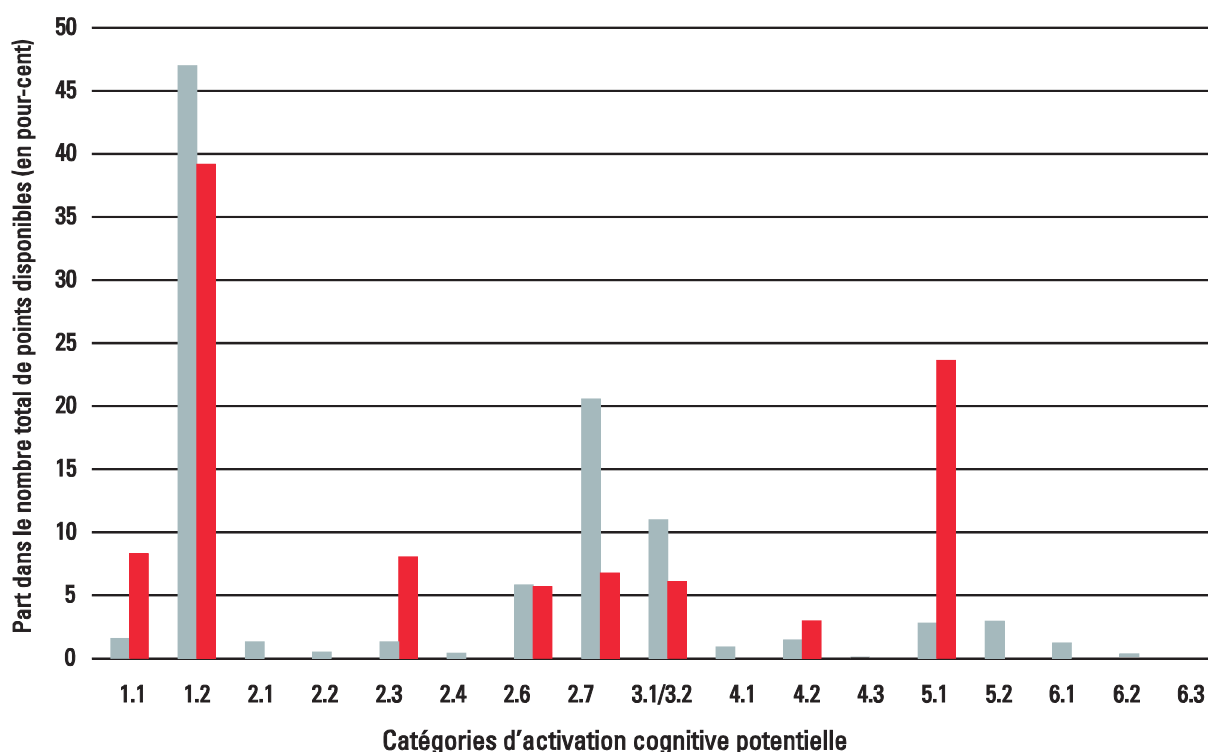
Catégorie principale et sous-catégories	Définitions adaptées pour la discipline de la biologie dans EVAMAR II <i>[d'après Anderson et al. (2001), p. 63–92]</i>	Part du nombre total de points pouvant être obtenus par la résolution des questions de cette catégorie w_{pH} en pour-cent
1	rappeler faire appel à un savoir pertinent de sa mémoire à long terme	
1.1	trouver dans sa mémoire à long terme un savoir en relation avec les données reçues	8.3%
1.2	faire appel à des connaissances pertinentes retenues par la mémoire à long terme	39.2%
2	comprendre saisir des contenus d'enseignement et en dégager du sens (oralement, par écrit, par une représentation graphique)	
2.1	transposer des données reçues en modifiant la forme de présentation	0.0%
2.2	trouver un exemple ou une situation imagée qui illustre une notion ou un principe	0.0%
2.3	décider de l'appartenance d'une instance à telle ou telle catégorie (de notions ou de principes)	8.1%
2.4	à partir de données reçues, trouver un sujet général ou résumer les éléments principaux (avec un processus personnel d'intégration) étudier les données reçues pour en tirer la quintessence, la pointe, «l'enjeu» et résumer (avec un processus personnel d'intégration)	0.0%
2.6	reconnaître les convergences et les différences entre des idées, des objets et autres ordonner et classer des idées, des objets et autres au sein d'un même niveau hiérarchique (anglais: mapping)	5.8%
2.7	appliquer à des éléments que l'on découvre un modèle, appris auparavant, décrivant un rapport de cause à effet appliquer signifie: si le donné = le ou les effet(s): utiliser le modèle pour remonter aux causes si le donné = la ou les cause(s): utiliser le modèle pour conclure aux effets	6.8%
3	appliquer appliquer ou utiliser une procédure dans une situation donnée	
3.2	appliquer une procédure (étapes qui s'enchaînent ou qui ont pu être apprises) à un problème <u>non</u> connu	6.2%
4	analyser diviser une donnée reçue en ses éléments constitutifs, définir les rapports mutuels de ceux-ci et la relation qu'ils entretiennent avec la structure ou l'objectif d'ensemble	

Catégorie principale et sous-catégories	Définitions adaptées pour la discipline de la biologie dans EVAMAR II <i>[d'après Anderson et al. (2001), p. 63–92]</i>	Part du nombre total de points pouvant être obtenus par la résolution des questions de cette catégorie w_{PH} en pour-cent
4.1	discerner des éléments pertinents (ou importants) et d'autres qui ne le sont pas au sein des données reçues (sans processus personnel d'intégration)	1.0%
4.2	déterminer comment des éléments prennent place et interagissent au sein d'une structure structurer des données reçues et les transposer sous une autre forme (attribuer des sous-titres, construire un tableau, élaborer un diagramme en bâtons, un histogramme, un topogramme, etc.)	2.9%
4.3	à partir des données reçues, reconnaître le point de vue d'un auteur, les distorsions de perception et de jugement, sa représentation des valeurs ou ses intentions	0.0%
5	évaluer émettre un jugement sur la base de critères et de standards	
5.1	détecter des illogismes ou des erreurs au sein d'une démarche ou d'un produit; décider si la démarche ou le produit est en soi logique; déterminer l'effectivité d'une procédure (telle qu'elle est appliquée)	22.5%
5.2	déceler, en fonction de critères externes, les points faibles d'un processus ou d'un produit; définir la cohérence externe d'un produit; reconnaître si une procédure est appropriée à un problème donné	0.0%
6	créer réunir des éléments pour former un ensemble fonctionnel cohérent; articuler des éléments selon une nouvelle structure	
6.1	élaborer des hypothèses	0.1%
6.2	élaborer un mode de faire pour résoudre un problème posé	0.0%
6.3	inventer un nouveau produit	0.0%

Les trois principaux résultats sont les suivants:

- a) Là encore, la moitié environ du nombre total de points disponibles dans un examen universitaire moyen en biologie, c'est-à-dire 47,6%, peut être obtenue par la résolution de simples questions de connaissances (catégories 1.1 et 1.2), c'est-à-dire de tâches qui ne requièrent pas de processus cognitif dépassant le rappel de savoir enregistré dans la mémoire.
- b) Comme pour les examens de maturité, la répartition par exigences cognitives est très inégale. Les étudiants et les étudiantes qui maîtrisent bien les quatre principaux processus cognitifs ci-après s'assurent déjà 78% du nombre total de points:
 - trouver dans sa mémoire à long terme un savoir en relation avec les données reçues (1.1),
 - faire appel à des connaissances pertinentes retenues par la mémoire à long terme (1.2),
 - décider de l'appartenance d'une instance à telle ou telle catégorie (de notions ou de principes)(2.3),
 - détecter des illogismes ou des erreurs au sein d'une démarche ou d'un produit; décider si la démarche ou le produit est en soi logique; déterminer l'effectivité d'une procédure (telle qu'elle est appliquée)(5.1).
- c) La part du nombre total de points qui peut être obtenue par la résolution des tâches de transfert et de métacognition se monte à 38,5%, soit quelques points de pourcentage de moins seulement que la valeur de 42,2% pour les examens de maturité (voir figure V.3).

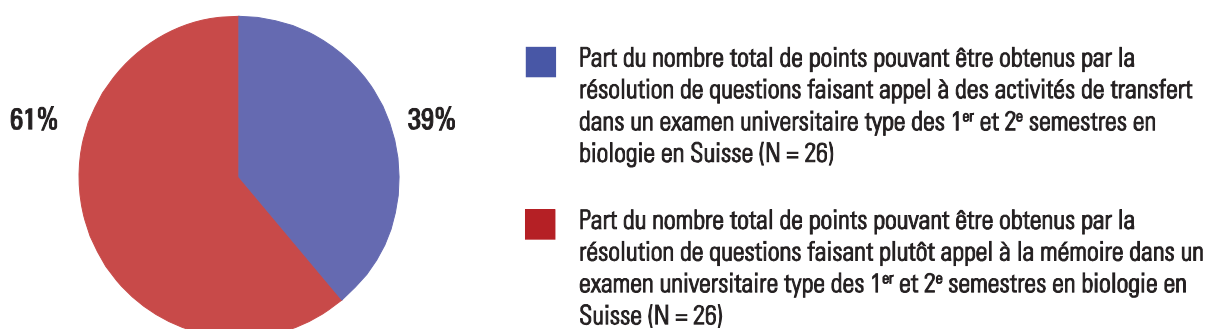
Figure V.3: Comparaison des exigences cognitives des examens de maturité et des examens universitaires en biologie



La comparaison des exigences cognitives potentielles entre les examens de maturité (en gris) et les examens universitaires (en rouge) montre par ailleurs que les questions de la catégorie 5.1 (reconnaître les incohérences et les erreurs) sont beaucoup plus fréquentes à l'université qu'au gymnase. A l'inverse, les tâches de la catégorie 2.7 (appliquer un modèle de cause à effet connu à une situation jamais rencontrée auparavant) se révèlent être une «spécialité» des examens de maturité.

La résolution de toutes les tâches faisant plutôt appel à la mémoire ne permet, seule, d'obtenir que 61,5% du nombre total de points disponibles dans un examen universitaire type en biologie. Le reste des points doit être obtenu en faisant appel à des activités de transfert (voir figure V.4).

Figure V.4: Parts respectives des tâches de mémoire et de transfert (biologie, université)



La comparaison entre les examens de maturité et les examens universitaires permet de tirer la conclusion générale suivante: les examens universitaires en biologie comportent certes beaucoup de questions à choix multiple, mais trois quarts de celles-ci ne peuvent pas être résolues par une simple activité de mémoire. Les options de ces tâches à choix multiple exigeantes sont constituées de phrases complètes dont il n'est possible de

juger la justesse ou le caractère erroné de chaque partie que par une compréhension générale de toute la phrase. Ces questions testent ainsi la qualité de la compréhension et peuvent donc être comptées comme des tâches à haut potentiel d'activation cognitive. Presque 40% du nombre total des points disponibles dans un examen universitaire en biologie doit donc être obtenu par la résolution de tâches qui ne peuvent pas être maîtrisées par une simple activité de mémoire. Ce pourcentage est presque aussi élevé que celui constaté pour les examens de maturité en Suisse en 2007 (42%). Ce résultat est très étonnant au vu des motifs exposés lors de la formulation de l'hypothèse. Visiblement, nombre de professeurs ont trouvé dans la formulation de questions à choix multiple complexes un moyen de concilier les exigences d'efficacité de la correction selon des critères objectifs et les exigences cognitives d'un examen approfondi de la compréhension. Effectivement, la recherche en matière de développement de tests corrige de plus en plus le préjugé selon lequel les questions à choix multiple ne se prêteraient qu'à tester les connaissances factuelles (pour une synthèse des travaux actuels, voir par ex. Haladyna, Downing & Rodriguez, 2002).

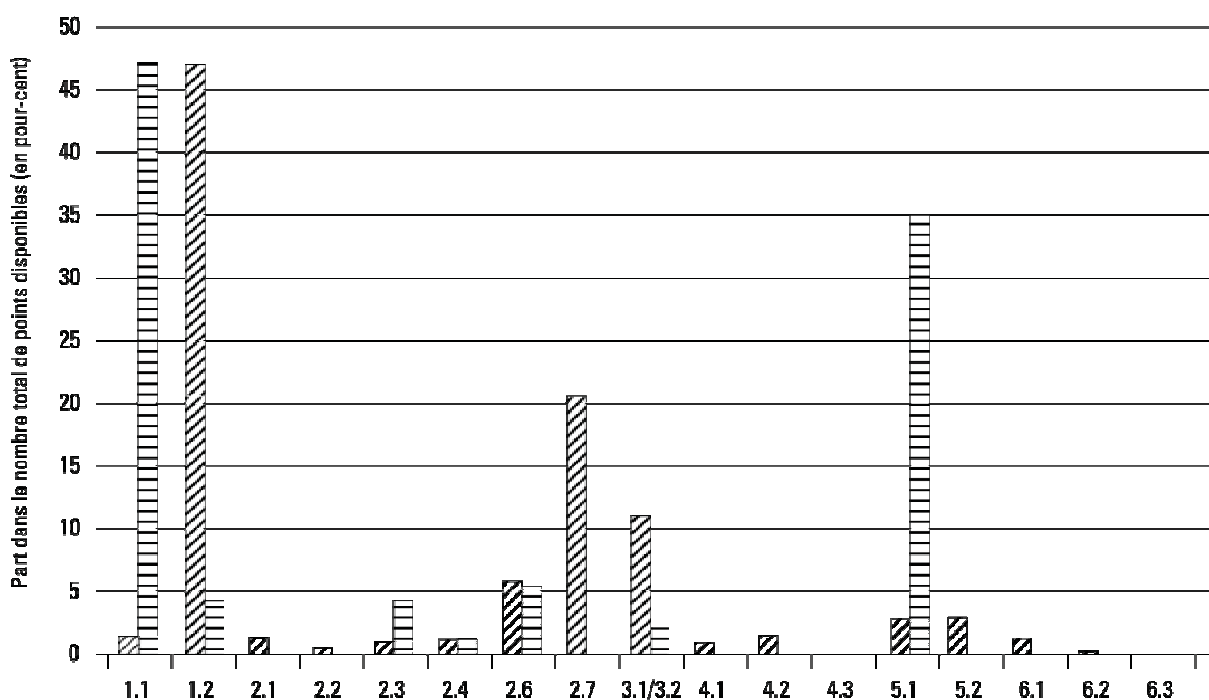
3.3.4 Comparaison avec les tâches du module C

Les items du test de biologie du module C avaient pour objet principal de vérifier dans quelle mesure les élèves disposaient des connaissances et capacités spécifiques en biologie nécessaires pour comprendre sans aide extérieure les contenus des supports de cours du premier semestre universitaire en biologie. La plupart des items sous forme de question à choix multiple étaient formulés de sorte à minimiser le travail de correction. A des fins de comparaison, nous avons estimé le potentiel d'activation cognitive des items sélectionnés pour le test de biologie dans EVAMAR II de la même manière que nous l'avons fait pour les questions des examens de maturité et des examens universitaires (voir tableau V.32).

Tableau V.32: Répartition des items sélectionnés pour le test de biologie dans EVAMAR II par catégories de potentiel d'activation cognitive

Catégorie	Format des items	Nombre d'items	Part w_{pE}
1.1	<i>Choix multiple</i> Avec options sous forme de mots-clés; les options reproduisent une définition donnée dans le cours, il est donc possible de reconnaître directement l'option correcte. La tâche est formulée comme une question, et non comme une affirmation.	43	47.3%
1.2	<i>Ouvert</i> Rappeler une notion spécialisée enregistrée dans la mémoire à long terme.	4	4.4%
2.3	<i>Choix multiple</i> Une seule option (instance ou principe) est correcte. Relier correctement des instances à des notions données.	4	4.4%
2.4	<i>Ouvert</i> A partir d'un diagramme donné, trouver le type de diagramme.	1	1.1%
2.6	<i>Choix multiple</i> Les options comparent deux notions, et seule une comparaison est correcte.	5	5.5%
3.1	<i>Choix multiple</i> Appliquer un schéma à de nouvelles illustrations, par ex. interpréter correctement un arbre phylogénétique. L'application correcte permet de choisir la bonne option.	2	2.2%
5.1	<i>Choix multiple</i> La tâche est formulée comme une affirmation tirée du cours plus une question. Quatre réponses sont données, dont trois options incohérentes avec l'affirmation et une option cohérente avec cette dernière. Pour juger de cette cohérence, une bonne compréhension de la biologie est nécessaire; il faut comprendre le sens des affirmations.	32	35.2%
Total des items et somme des pourcentages		91	100.1%

Figure V.5: Comparaison de la répartition par potentiels d'activation cognitive des tâches des examens de maturité 2007 en Suisse (hachures obliques) et des items sélectionnés pour l'évaluation EVAMAR II en biologie (hachures horizontales)



La figure V.5 compare les questions de maturité avec celles du test de biologie. On constate d'emblée que les items du test de biologie d'EVAMAR II n'appellent pas d'activité de mémoire pure dans un format ouvert, mais sous la forme de question à choix multiple qui comporte des options sous forme de mots clé, parmi lesquelles il suffit de reconnaître la bonne réponse. Dans les examens de maturité, la mémoire est beaucoup plus souvent sollicitée sous une forme active. A ce propos, il convient de rappeler qu'en regard des quelque 3800 élèves soumis au test, l'efficacité de la correction était nécessairement un critère prioritaire. On observe par ailleurs que dans les examens de maturité, les tâches destinées à examiner le degré de compréhension prennent souvent la forme d'une tâche dotée d'un potentiel d'activation cognitive de la catégorie 2.7. Etant donné que les tâches de ce type conduisent généralement à des réponses de plusieurs phrases difficiles à corriger, les concepteurs des items du test leur ont préféré une solution analogue à celle apparemment privilégiée à l'université. En effet, un grand nombre d'items de compréhension du test de biologie dans EVAMAR II prennent la forme de question à choix multiple qui activent potentiellement le processus cognitif exigeant de la reconnaissance des options erronées.

3.3.5 Comparaison avec les résultats du sous-projet A4

Le tableau V.33 résume les déclarations des professeurs des disciplines de la biologie et de la médecine concernant les déficits de connaissances et de capacités constatés chez les étudiants débutant en matière de biologie. Pas moins de 6 professeurs de biologie et de médecine sur 21 ayant retourné le questionnaire expriment dans leurs réponses le souhait que les étudiants aient de meilleures connaissances dans les sous-domaines de l'écologie et de la systématique/évolution. Seuls 2 professeurs sur 21 pointent des carences dans d'autres sous-domaines. Les professeurs de biologie et de médecine situent donc les plus grands déficits de leurs élèves précisément dans les sous-domaines qui ont relativement le moins de poids dans les examens de maturité en langues française et allemande en Suisse. Précisons que la répartition thématique des questions d'examen ne dit encore rien sur la répartition thématique du temps d'enseignement ni sur la difficulté de ces sous-domaines, ces deux points n'ayant pas été examinés.

Tableau V.33: Déficiences constatées en biologie et savoir et capacités supplémentaires souhaités chez les nouveaux étudiants
(en italique: écologie et systématique/évolution)

Identifiant du professeur	Question 3: Déficiences de savoir et de capacités constatées actuellement	Question additionnelle 6: Savoir et capacités supplémentaires souhaités chez les nouveaux étudiants
23	aucune idée de ce qu'est la photosynthèse connaissance très fragmentaire des processus de la mitose et de la méiose <i>déficiences, voire souvent absence totale de connaissances de base en écologie</i>	<i>traiter plus de questions et de problèmes liés à l'écologie</i>
39	<i>de façon générale, la compréhension et le savoir en matière de biologie de l'évolution semblent être les moins bien partagés</i>	
55	compréhension des mécanismes chimiques de la biologie	
78	<i>expérience directe dans la nature (grâce à des excursions, des travaux de semestre ou de maturité sur site, etc.)</i> connaissance des formes connaissance de la crise actuelle de la biodiversité et des principales raisons de la disparition des espèces en Europe en partie aussi compréhension des arbres phylogénétiques	<i>expérience directe dans la nature (grâce à des excursions, des travaux de semestre ou de maturité sur site, etc.)</i> connaissance des formes connaissance de la crise actuelle de la biodiversité et des principales raisons de la disparition des espèces en Europe
106		<i>Ecologie: connaissance des problèmes écologiques de la Suisse et du monde</i>
51		L'enseignement de la biologie ne doit pas transmettre en premier lieu des savoirs factuels en matière de biologie cellulaire et moléculaire moderne, mais plutôt la connaissance des cycles des éléments, des interdépendances écologiques, de la diversité des espèces et de l'évolution.

3.3.6 Evaluation globale des examens de maturité en biologie en Suisse

A notre connaissance, il n'existe dans la littérature qu'une seule autre étude qui examine l'importance relative de chaque type de tâches en termes d'incidence sur le nombre total de points disponibles en se référant à des examens de biologie. L'étude situe le pourcentage du total de points disponibles déterminés par les «higher-order questions» (catégories principales 3 à 6 selon Bloom) entre 21% et 52% (Zheng et al., 2008). Toutefois, les examens étudiés ne sont pas destinés à être passés au sein d'une classe, puisqu'il s'agit de tests de grande envergure largement diffusés aux Etats-Unis (le *Medical College Admission Test* MCAT, l'*Advanced Placement AP Biology Course*, l'*introductory biology course for undergraduate majors*) de trois universités différentes, du *Biology Graduate Record Examination* GRE et de plusieurs «firstyear medical school courses from an institution with a traditional curriculum»). Il est réjouissant de constater que les examens de maturité en biologie organisés en Suisse, en affichant une valeur correspondante de 40,8%, se situent tout à fait dans la moyenne de ces gros tests, élaborés avec des moyens importants. Ce résultat ne va pas de soi, et ce pour trois raisons:

- Aux Etats-Unis, les milieux pédagogiques ont commencé dès les années 1960 à promouvoir de façon ciblée les «capacités intellectuelles d'ordre supérieur» («higher-order thinking»), y compris en travaillant sur le type des questions d'examen. Chipman, Segal & Glaser (1985) et Resnick (1987) font le bilan (parfois emprunt de scepticisme) des différents programmes de formation lancés à ce titre, comme le «productive thinking program», «teaching thinking» ou encore «improving intelligence». Quarante ans plus tard, la proportion importante de «higher-order questions» dans les tests actuels pratiqués aux Etats-Unis n'a donc rien d'étonnant. Dans les pays germanophones, ce n'est que dix à vingt ans plus tard que les premières tentatives pour la promotion de la pensée productive (Kretschmer, 1981) ou les

«entraînements de la pensée» («Denktrainings») ont vu le jour (voir à ce propos les aperçus rétrospectifs de Friedrich & Mandl, 1992, et Hager, 1995). Si l'on en croit l'expérience, il faut des années, voire des décennies pour que des réflexions universitaires assurées sur un plan empirique se répercutent concrètement sur la formation des enseignants. Il était donc tout à fait plausible que la formation systématique de nouveaux enseignants dans la rédaction de tâches d'examen exigeantes ait débuté il y a quelques années seulement dans l'espace germanophone et ne montre donc pas encore de résultats probants. Les résultats d'EVAMAR II contredisent toutefois cette hypothèse en montrant que cette évolution a manifestement déjà commencé plus tôt et porte déjà ses fruits.

- b) L'élaboration de bonnes tâches de transfert nécessite beaucoup de temps et une très grande expertise scientifique, que seule une recherche menée spécialement pour la question d'examen permet d'acquérir dans la plupart des cas. Manifestement, beaucoup d'enseignants sont convaincus que le jeu en vaut la chandelle.
- c) La correction de bonnes tâches de transfert demande un grand investissement intellectuel et de temps. Cela place sans doute souvent les enseignants dans une situation paradoxale: selon l'étude de Bol & Strage (1996), les enseignants de biologie tendent à limiter les questions de transfert dans leurs examens, bien qu'ils considèrent le développement des capacités de transfert comme l'un des buts de leur enseignement. Là encore, la conviction que l'enjeu justifie l'effort semble l'emporter.

La proportion élevée de tâches cognitivement exigeantes dément également le vieux cliché prétendant qu'il suffirait d'apprendre «par cœur» sans réfléchir pour obtenir facilement une bonne note en biologie. Les résultats suggèrent cependant également une possible négligence des domaines thématiques de l'écologie et de la systématique/de l'évolution. Les indications des professeurs et le poids relativement faible de ces thèmes dans la note de l'examen en biologie peuvent être interprétés comme un indice que ces matières sont traitées de façon trop superficielle dans la perspective de la capacité de faire des études et qu'elles devraient recevoir une place plus importante dans les examens.

Les bacheliers et les bachelières ne sont pas confrontés à des tâches d'un degré de difficulté jamais rencontré lorsqu'ils commencent des études. C'est la conclusion qui ressort de la comparaison entre les potentiels d'activation cognitive des examens de maturité, par définition connus des bacheliers, et ceux des examens de première année à l'université. Au vu du vaste savoir factuel sur lequel se fonde la biologie, une bonne mémoire, telle qu'elle est exercée par «l'apprentissage par cœur» dans l'enseignement de la biologie au gymnase, reste une base solide pour réussir également aux examens universitaires (cf. les pourcentages de tâches de la catégorie 1.2 à l'université et au gymnase, respectivement de 39% et de 47%). La plus grande différence entre le gymnase et l'université concerne la catégorie 5.1, qui se réfère aux tâches exigeant de reconnaître précisément les erreurs glissées dans des affirmations spécifiques. Les questions à choix multiple avec des options à qualifier de justes ou de fausses sont étonnamment difficiles pour beaucoup d'étudiants, ce qui s'explique par la difficulté réelle qu'il y a à repérer des incohérences parfois subtiles dans des affirmations scientifiques formulées avec une grande densité. Néanmoins, il serait peut-être possible d'améliorer encore l'aptitude aux études de nos bacheliers et de nos bachelières en exerçant de façon ciblée ce type de questions et en l'utilisant dans les examens de maturité.

Sous-projet D2: la qualité des travaux de maturité en Suisse

Ce sous-projet a été entièrement conçu et réalisé par les auteurs suivants de la Fachhochschule Nordwestschweiz: Christina Huber, Vera Husfeldt, Lukas Lehmann, Carsten Quesel (direction, carsten.quesel@fhnw.ch). L'équipe a bénéficié du soutien de Felix Boller (PH FHNW) et d'Eva Roos (IRD).

Introduction

L'évaluation de la qualité des travaux de maturité, qui fait l'objet du présent rapport, doit en particulier porter sur leur valeur formatrice en termes de qualifications transversales, d'interdisciplinarité et de propédeutique scientifique, à la lumière des objectifs généraux de formation définis dans le RPM 95 (cf. Eberle, 2005). Elle doit aussi être affinée dans le prolongement des résultats de la première phase d'évaluation (EVAMAR I). Elle examine notamment les facteurs qui influent sur la qualité des travaux de maturité. Le sous-projet D2 doit d'abord déterminer les traits distinctifs de la qualité d'un travail de maturité. Ce n'est qu'une fois formulé un tel catalogue de critères qu'il devient possible d'évaluer concrètement des travaux d'élèves, puis de porter une appréciation sur la qualité des travaux de maturité en Suisse.

Outre l'analyse des travaux de maturité, il convient aussi d'observer plus précisément le cadre institutionnel général mis en place par les cantons et les gymnases, car il a un effet déterminant sur les objectifs du travail de maturité et ses méthodes de préparation. A un troisième niveau, l'analyse doit envisager les élèves eux-mêmes; ils ont été directement interrogés sur la préparation de leur travail de maturité, mais aussi sur les effets de ce processus: choix du sujet, encadrement et assistance, ressources, appréciation du volume de travail investi et utilité du travail de maturité surtout. Enfin, l'étude examine les facteurs qui, au niveau de l'école ou de l'élève, ont un impact sur la qualité du travail de maturité.

Dans cette optique, le présent projet comporte trois volets. Tout d'abord, 920 élèves ont été consultés sur la base d'un questionnaire standard écrit, et 437 travaux de maturité au total ont fait chacun l'objet de deux évaluations indépendantes fondées sur une grille de notation standardisée prédéfinie. Ces deux composantes de la recherche sont quantitatives. Pour le troisième volet, l'analyse des cadres mis en place par les cantons et les écoles, on a eu recours à une analyse qualitative, à laquelle ont participé 30 gymnases de toute la Suisse.

Pour diverses raisons, cette étude a eu ses limites. Il n'a par exemple pas été possible d'appliquer une méthodologie spécifique aux avis des enseignants chargés de suivre les travaux de maturité. L'enquête n'étant pas réalisée en parallèle avec le processus, il a en outre été impossible d'analyser les interactions pédagogiques auxquelles ont donné lieu les travaux de maturité, et de mesurer l'acquisition des compétences en cours de travail chez les élèves. Enfin, il n'a pas pu être procédé à la validation des pratiques d'évaluation dans les diverses régions linguistiques en ce qui concerne l'uniformité des normes, comme cela avait été initialement prévu, faute de disposer d'un nombre suffisant d'évaluateurs possédant les compétences linguistiques et spécialisées nécessaires.

Structure du présent chapitre

Le présent chapitre présente la réflexion théorique et les principales notions à la base de l'étude, ainsi que l'état actuel de la recherche empirique sur les travaux de maturité (section 1). Sont ensuite décrites la conception de l'enquête et les méthodes concrètement mises en œuvre (section 2).

Les résultats empiriques sont présentés en seconde moitié de chapitre: la section 3 est consacrée aux cadres mis en place par les cantons et les écoles, la section 4 aux résultats de l'enquête menée auprès des élèves. Enfin, la section 5 présente et analyse les résultats de la notation.

Remerciements

Nous tenons à remercier en première place les élèves qui ont pris le temps de remplir nos questionnaires et ont mis leurs travaux de maturité à notre disposition pour évaluation. Nous remercions également les écoles qui ont participé à l'étude, et en particulier les personnes en leur sein qui nous ont fourni renseignements et documentation.

Nos remerciements vont aussi aux experts qui ont accepté de se former à la manipulation des grilles de critères, puis d'évaluer les travaux de maturité.

Dans le canton du Tessin, nous sommes reconnaissants de leur aide à MM. Emanuele Berger (directeur de l'Ufficio studi e ricerca, USR) et Stefano Chiantese pour la traduction. En Suisse francophone, notre reconnaissance va à nos partenaires de l'Institut de recherche et de documentation pédagogique (IRDp) de Neuchâtel, et notamment à M. Matthis Behrens, son directeur, et à Mme Eva Roos, collaboratrice scientifique, qui nous ont aidés à mettre au point les instruments et à collecter les données en Suisse romande.

M. Felix Boller (PH FHNW) a beaucoup contribué à la préparation de la panoplie d'instruments; c'est pour une part déterminante à sa longue expérience de l'enseignement gymnasial et de la didactique que l'étude doit son caractère pratique, ce dont nous le remercions spécifiquement ici.

1 Contexte général et bases théoriques

1.1 Le travail de maturité dans la réforme de 1995

1.1.1 Introduction du travail de maturité

La genèse de l'actuel travail de maturité remonte à plusieurs décennies. Le rapport *L'Enseignement secondaire de demain* de la commission mandatée par la CDIP avait proposé en 1969 l'introduction de travaux écrits (Meylan, 1996, p. 19 s), faisant valoir que l'introduction de «travaux personnels de plus grande envergure», faisant partie intégrante de l'examen de maturité, permettrait d'approfondir l'apprentissage gymnasial et donnerait l'occasion aux futurs étudiants de «prendre conscience de leurs dispositions individuelles» (CDIP, 1972, p. 105). Ces travaux prouveraient «que le futur bachelier est capable de penser de manière originale, qu'il sait dégager l'essentiel d'une situation et qu'il est doté d'un esprit logique et critique»; en outre, ils permettraient au candidat de «montrer qu'il possède une imagination créatrice et qu'il connaît les méthodes essentielles conduisant à la solution des problèmes posés» (CDIP, 1972, p. 107). Enfin, ces travaux personnels donneraient «une forme plus individualisée» aux examens de maturité (CDIP, 1972, p. 95). Le projet prévoyait que les élèves préparent un travail autonome dans deux ou trois matières, et que l'examen de maturité comporte pour chacun d'eux une épreuve revêtant la forme d'un entretien.

Mais l'évaluation du projet gymnasial n'est devenue juridiquement obligatoire qu'à la suite de la révision de l'ordonnance fédérale sur la reconnaissance des certificats de maturité (ORM), dans les années 90. L'art. 10 du règlement de 1995 sur la reconnaissance des certificats de maturité gymnasiale (RRM) stipule que chaque élève doit effectuer, seul ou en équipe, un travail autonome d'une certaine importance, qui fait ensuite l'objet d'un texte ou d'un commentaire rédigé et d'une présentation orale.

1.1.2 Interdisciplinarité du travail de maturité

A l'origine de l'introduction du travail de maturité, il y avait surtout l'intention de renforcer l'apprentissage interdisciplinaire et de développer des compétences clés cognitives, personnelles et sociales (CDIP, 2000, p. 115 s; Meylan, 1996, p. 31 s). Nul ne conteste aujourd'hui que le travail de maturité remplit une double fonction d'enrichissement et de démonstration des compétences en fin de formation gymnasiale. Il est même parfois considéré comme la seule authentique innovation du RRM (Dreyer, 2005, p. 23), ou l'un des indéniables points forts de la réforme (Hirschi, 2002, p. 29). Les enseignants des gymnases confirment constamment qu'il a

donné les preuves de sa valeur pédagogique (cf. Bonati, 2006, p. 6; Jacob, 2004, p. 5). Mais de nombreux signes montrent que le potentiel d'interdisciplinarité du travail de maturité est loin d'être épuisé. Certains critiques observent que dans la plupart des cas, il se limite à une seule matière (cf. Schoch, 2003; Kunz, 2003). Les élèves interrogés dans l'enquête EVAMAR I n'avaient d'ailleurs choisi que dans 24% des cas un sujet pluridisciplinaire (Ramseier et al., 2005, p. 11).

Les décisions politiques et administratives de la DDIP et de la Confédération témoignent de leur volonté de revaloriser le travail interdisciplinaire à l'avenir. A la faveur de la dernière révision du RRM, entrée en vigueur pour l'année 2008-2009, il a été explicitement précisé à l'article 11^{bis} que «chaque école pourvoit à ce que les élèves soient familiarisés aux approches interdisciplinaires».

1.1.3 Le travail de maturité dans l'examen de maturité

Dans tous les cantons sauf celui de Lucerne⁴⁴, le travail de maturité ne figure pas pour l'instant directement dans les épreuves de maturité telles que les définit l'article 16 du RRM. Le certificat de maturité n'en comporte pas moins l'appréciation des épreuves écrites et orales liées au travail de maturité, avec mention du sujet (art. 20 RRM); et dans la plupart des cantons, l'élève n'est admis à se présenter aux épreuves de maturité que moyennant une note suffisante au travail de maturité. En cas de succès, le titre et la note de travail de maturité donnent à l'élève une sorte de carte de visite qui peut lui inspirer plus ou moins de fierté (Fontolliet, 1996, p. 55).

Cela dit, les élèves se montrent souvent critiques à l'égard du poids insuffisant à leurs yeux du travail de maturité dans l'obtention du certificat (cf. Binder & Feller-Länzlinger, 2004, p. 52; 2003, p. 22; Davaud & Hexel, 2003, p. 67; USO, 2006); 62% des élèves interrogés dans le cadre de l'étude EVAMAR I trouvaient que le travail de maturité ne comptait pas suffisamment dans l'obtention de la maturité. On observait à ce niveau de sensibles variations régionales: une petite moitié seulement d'entre eux étaient de cet avis en Suisse francophone, contre les deux tiers en Suisse germanophone (63,4%) et italophone (66,1%; cf. Pagnossin, Alliaat & Dozio, 2005, p. 190). Mais le rapport de la première étude EVAMAR n'expliquait pas le phénomène.

En revanche, une enquête menée auprès d'enseignants de gymnases argoviens révèle qu'ils seraient plutôt favorables à l'abandon de la comptabilisation du travail de maturité dans l'examen, même s'ils comprennent le désir des élèves. Ils justifient leur position en expliquant qu'un coefficient plus élevé accordé au travail de maturité exigerait une notation fondée sur des normes impératives et comparables. La multiplicité infinie des sujets et la diversité de la façon dont ils sont traités rendent toutefois cette standardisation difficile. Certains enseignants observent par ailleurs que l'absence de peur d'une mauvaise note renforce la motivation intérieure des élèves et fait contrepoids à la pression dominante des notes, c'est-à-dire à la motivation extérieure au travail (Binder & Feller-Länzlinger, 2004, p. 51 s).

Même si la question de la standardisation des procédures d'évaluation et des conditions générales de réalisation des travaux de maturité reste toujours sans réponse, la révision partielle du RRM lancée en 2007 stipule que le travail de maturité est désormais pris en compte «sur la base de la mise en œuvre du projet, du document déposé et de la présentation orale» (cf. CDIP, 2007; Oelkers, 2008, p. 203 ss).

1.2 Le travail de maturité dans la formation gymnasiale

La formation gymnasiale concentrait traditionnellement ses programmes sur les langues anciennes et débouchait fonctionnellement sur une université plaçant la théologie, la médecine et le droit au cœur des disciplines académiques. Au 19^e siècle, les langues anciennes ont subi au gymnase la pression des réformes d'inspiration

⁴⁴ Dans le règlement des examens de maturité du canton de Lucerne (du 27 mai 1999), le travail de maturité figure parmi les matières indispensables à l'obtention du diplôme.

économique et technique qui revalorisaient notamment les sciences de la nature. Ce qui a conduit dans de nombreux pays à un clivage entre les établissements à vocation littéraire ou scientifique et technique, voire plus tard économique et sociale. Cette différenciation a inévitablement soulevé la question des critères d'admission aux études supérieures. Malgré de vives discussions sur la création de maturités spécialisées, on en est finalement resté à la maturité générale, point d'aboutissement ultime du cycle secondaire⁴⁵.

La différenciation des sciences, au cours du 20^e siècle, a encore multiplié les ambiguïtés en ce qui concerne la fonction de propédeutique scientifique des études gymnasiales. Si l'on estimait jusque-là relativement naturel de suivre les lignes de démarcation du savoir présentées dans les manuels, et de définir sur cette base un spectre relativement clair de disciplines, il a fallu repenser le découpage à la lumière de la spécialisation et de la diversification de l'activité universitaire. Ce qui a conduit au cours de ces dernières décennies à réévaluer le modèle de culture générale qui sous-tendait les programmes d'enseignement gymnasial. Cela apparaît clairement dans le plan d'études cadre des écoles de maturité (CDIP, 1994), qui prévoit que le gymnase ne peut ni ne doit dispenser une spécialisation propédeutique pour certaines filières, mais fournir aux élèves une large base de connaissances, de savoir-faire et d'attitudes pour les préparer à enrichir leurs connaissances dans n'importe quel domaine et à tout moment (CDIP, 1994, p. 6; cf. Oelkers 2008, p. 74 ss).

L'idée de l'apprentissage autodirigé a été lancée dès le début du 20^e siècle, dans le contexte de la pédagogie du projet développée dans le sillage de l'éducation progressive. Dans la pédagogie du projet, c'est l'enfant lui-même qui conçoit l'enseignement, ce se sont sa curiosité et son activité personnelle qui priment, et l'apprentissage consiste surtout pour lui à s'attaquer activement à des problèmes et à chercher à les résoudre par lui-même ou avec d'autres (Frey, 1993; Oelkers, 1997). Dans ce nouveau courant pédagogique, la recherche et l'expérimentation avaient leur place dès l'école primaire. La collecte et l'analyse indépendantes d'informations étaient considérées comme la base de l'apprentissage scolaire; les études scientifiques universitaires ne faisaient que prolonger les stratégies cognitives acquises dès l'enfance.

Tout d'abord perçue comme un modèle d'organisation de l'école, la pédagogie du projet est de plus en plus devenue au cours du 20^e siècle un moment auxiliaire dans un programme d'études largement différenciées selon les critères didactiques des diverses disciplines. Ce redimensionnement du rôle du projet reflète surtout l'idée qu'un centrage radical sur l'enfant déboucherait en fin de compte sur une non-directivité pédagogique totale pour ce qui est de l'organisation et de la structuration des processus d'apprentissage (Oelkers, 1997, p. 27 s). C'est pourquoi le travail de maturité est aussi un moment complémentaire dans le programme gymnasial. Le large spectre de ses exigences en fait, au-delà de l'enseignement dispensé dans les disciplines établies, un élément important d'accomplissement de la mission de formation des gymnases (cf. Binder, 2003; Eberle, 2007; Hadorn, 2004).

Cette mission de formation présente aujourd'hui trois volets. La Conférence permanente des ministres de l'éducation et des affaires culturelles des länder allemands (KMK) définit ainsi les trois objectifs de formation gymnasiale dans le concordat sur le degré supérieur du second cycle de formation secondaire: l'enseignement doit y donner aux élèves une culture générale approfondie, l'aptitude générale de poursuivre des études universitaires, ainsi qu'une formation propédeutique en sciences (KMK, 2006, p. 5). La mention de l'approfondissement de la culture générale assure la continuité avec les objectifs de formation du primaire et du premier cycle secondaire. Mais par ailleurs, la fonction de propédeutique scientifique s'appuie justement sur l'enrichissement constant de la culture générale, qui va au-delà de la méthodologie stricte et des méthodes formelles de travail des disciplines scientifiques, établit un lien avec la vie, et devient ainsi un vecteur transdisciplinaire de compréhension (von Hentig, 1996, p. 177). Il est intéressant de remarquer que la KMK fait ici la différence entre l'aptitude aux études supérieures et la formation propédeutique scientifique. Cette distinction est ensuite précisée en ce qui concerne l'enseignement: il doit non seulement familiariser par exemple les élèves avec les questions, les catégories et les méthodes scientifiques, mais aussi leur faire maîtriser

⁴⁵ Cf. Oelkers (2008, p. 185 ss) pour une comparaison internationale, et Schäfer-Koch, (1997, p. 197 s.) pour l'Allemagne.

une base de connaissances spécialisées (KMK, 2006, p. 5), ce qui fonde en grande partie la nécessité de lier des savoirs de divers domaines, et de favoriser des méthodes de travail indépendantes et systématiques. Cette base de connaissances est elle-même perçue comme nécessaire au développement de la capacité de se prendre en charge, de l'aptitude au travail d'équipe et des capacités de communication.

Cette triple mission qu'assigne la KMK aux gymnases allemands se retrouve à l'art. 5 du RRM consacré aux objectifs de formation: dispenser «une formation générale équilibrée et cohérente, qui confère aux élèves la maturité requise pour entreprendre des études supérieures et les prépare à assumer des responsabilités au sein de la société actuelle» (RRM 95, art. 5, al. 1). Il va donc falloir maintenant définir plus précisément les notions de maturité ou d'aptitude aux études supérieures, de propédeutique scientifique et de compétences transversales que contient cette mission de formation formulée de façon relativement large.

1.2.1 Aptitude aux études supérieures et maturité intellectuelle

Pour Andrä Wolter, la notion d'aptitude aux études supérieures cache un insoluble problème multidimensionnel d'hétérogénéité (Wolter, 1997, p. 49), qui trouve son origine dans les processus de différenciation du système scientifique évoqué plus haut, mais aussi dans l'évolution générale de la société (multiplication des canaux d'information et de communication, des univers de vie quotidienne et des champs de socialisation). Ces changements se sont traduits par une hétérogénéité grandissante, aussi bien en ce qui concerne les conditions requises pour la formation et les attentes des étudiants que pour ce qui est des exigences des études supérieures et des attentes des universités (Wolter, 1997, p. 49). Des études empiriques montrent clairement que les étudiants de première année doivent répondre à des attentes très différentes selon la discipline, surtout pour ce qui est de leurs connaissances spécifiques et méthodologiques (cf. Konegen-Grenier, 2002, p. 166; Heldmann, 1998). Les opinions ne convergent que sur le fait que les diverses aptitudes relèvent de plusieurs dimensions: personnelle, formelle et matérielle (Konegen-Grenier, 2002); Konegen-Grenier (2002, p. 79) en vient à considérer l'aptitude aux études supérieures comme une faculté complexe, présentant des dimensions personnelles, sociales et techniques. Son enquête auprès d'enseignants universitaires la conduit à définir ainsi le contenu de l'aptitude aux études supérieures:

Les capacités d'analyse, d'abstraction et de différenciation sont indispensables dans les études supérieures. Elles doivent se combiner à de bonnes connaissances d'anglais, de mathématiques et d'allemand. La connaissance de la discipline doit être complétée par certaines techniques de travail: présentation, traitement de texte, méthodologie de la recherche. Pour appliquer ses capacités cognitives et son savoir, l'étudiant de première année doit avoir à l'égard de son travail une attitude personnelle alliant intérêt pour le fond et goût du travail et de la précision.

(Konegen-Grenier, 2002, p. 168; traduit de l'allemand).

Cette description du contenu de l'aptitude aux études supérieures reflète une association étroite avec les contenus de la propédeutique scientifique et des compétences transversales, mais sans les en distinguer suffisamment.

Dans l'enquête EVAMAR I, il avait été demandé aux élèves d'apprécier la préparation qu'ils avaient reçue aux études supérieures. Les réponses avaient été dans l'ensemble favorables puisque 76% d'entre eux se sentaient bien ou plutôt bien préparés (Ramseier, Allraum & Stalder, 2005, p. 127). Mais le tableau était toutefois contrasté d'une discipline à l'autre: la proportion variait entre 80% (pour l'histoire) et 14% (pour les sciences de l'ingénieur). Outre l'histoire, c'était surtout dans les langues, les sciences de la nature et en mathématiques que les étudiants estimaient avoir été bien ou plutôt bien préparés. Si la préparation avait été jugée la plus faible en sciences de l'ingénieur, c'est surtout parce qu'il s'agit d'un domaine qui occupe à bien des égards une position transversale dans le découpage des matières scolaires, si bien que les futurs étudiants ne peuvent pas s'appuyer sur leur pratique scolaire pour deviner à quoi s'attendre (Ramseier, Allraum & Stalder, 2005, p. 128). EVAMAR I avait par ailleurs montré que la préparation à la discipline choisie était en fait bonne: 54% des titulaires de la maturité avaient réagi par «plutôt d'accord», et 22% par «tout à fait d'accord» à l'affirmation «je suis bien préparé aux études que j'ai choisies» (Ramseier, Allraum & Stalder, 2005, p. 133).

TOSCA donne des résultats un peu moins flatteurs sur la durée pour les bacheliers du Bade-Wurtemberg qui, après avoir jugé bonne la qualité de la préparation aux études supérieures qu'ils avaient reçue au gymnase, se montraient rétrospectivement plus critiques une fois qu'ils avaient démarré dans leurs études supérieures. En moyenne, s'ils ne condamnaient pas leur école, ils ne se montraient pas vraiment confiants non plus dans la valeur du diplôme de maturité qu'ils y avaient acquis (Trautwein et al., 2006, p. 56 s).

1.2.2 Propédeutique scientifique

Le RRM ne contient que deux brèves mentions de la fonction de propédeutique scientifique du gymnase: l'art. 5 prévoit parmi les objectifs de formation que ces études doivent conférer aux élèves «la maturité requise pour entreprendre des études supérieures», et les familiariser «avec la méthodologie scientifique». Ainsi apparaissent les notions très proches de maturité nécessaire et d'aptitude aux études supérieures. D'un autre côté, la formulation indique clairement qu'il n'est guère possible de se fonder une pédagogie pour «enseigner» la propédeutique scientifique, qui renvoie davantage actuellement à des méthodes de travail et moins à des contenus (cf. Criblez, 2005, p. 5). Ce qui explique qu'elle soit aussi surtout considérée d'un point de vue méthodologique dans l'étude TOSCA menée dans le Bade-Wurtemberg; en relèvent par exemple la recherche documentaire, la rédaction de comptes rendus et la réalisation d'expériences (Trautwein & Lüdtke, 2004, p. 362 s).

Au niveau méthodologique, l'accent est nettement mis sur la notion de «littératie scientifique», très utilisée dans le monde anglo-saxon. La National Academy of Science américaine donne la définition suivante de cette compétence de base centrée sur la compréhension des règles de la logique scientifique et technique:

«Scientific literacy means that a person can ask, find, or determine answers to questions derived from curiosity about everyday experiences. It means that a person has the ability to describe, explain, and predict natural phenomena. Scientific literacy entails being able to read with understanding articles about science in the popular press and to engage in social conversation about the validity of the conclusions. Scientific literacy implies that a person can identify scientific issues underlying national and local decisions and express positions that are scientifically and technologically informed. A literate citizen should be able to evaluate the quality of scientific information on the basis of its source and the methods used to generate it. Scientific literacy also implies the capacity to pose and evaluate arguments based on evidence and to apply conclusions from such arguments appropriately.»

(National Academy of Science, 1996)

Si l'on reformule cette définition en allant au-delà de la perspective de la logique des sciences de la nature, on arrive à une conception de l'orientation et de la propédeutique scientifiques se référant à la capacité spécifique des élèves à maîtriser et à comprendre des modes opératoires spécifiques⁴⁶. Cette conception non liée à la logique de telle ou telle discipline est très proche de la *Science Literacy* considérée par Bonati comme une compétence gymnasiale caractéristique, et qui fournit une base permettant d'aborder des éléments essentiels de la méthodologie scientifique sous des formes simplifiées (Bonati, 2007, p. 51). Dans cette optique, l'orientation et la propédeutique scientifiques reviennent à ce que l'élève apprenne à appliquer des techniques élémentaires de production méthodiquement contrôlée de savoirs, et à dégager des indices qui le conduiront systématiquement à un jugement.

La différenciation du système scientifique et la spécialisation qui en découle font qu'il est difficile de définir un canon de règles de base valable pour toutes les disciplines. On peut toutefois esquisser les exigences générales auxquelles doit répondre une formation à caractère propédeutique scientifique. Selon Schmidt (1994, p. 266 ss), la propédeutique scientifique consiste surtout en une sensibilisation et une prise de conscience des élèves d'ordre méthodologique, d'autant plus que les savoirs scientifiques sont toujours obtenus sur une base méthodologique. Pour la formation gymnasiale, cela signifie que les élèves acquièrent des connaissances sur les grandes structures et méthodes des diverses sciences, c'est-à-dire se familiarisent avec les procédures

⁴⁶ Par opposition à la *Academic Literacy*, qui se limite largement au développement des compétences de lecture et d'écriture (cf. Neeley, 2005).

scientifiques et apprennent ainsi à les appliquer dans leur réflexion. Ils doivent aussi développer leur capacité et leur volonté de réfléchir sur leurs propres méthodes en termes de théorie du savoir et des sciences, de les comparer avec d'autres approches et de les replacer dans certaines traditions de recherche. Ils doivent pouvoir reconnaître les limites de la science, et comprendre que la production de connaissances, toujours liée à des intérêts, est soumise au changement et évolutive. Et il faut enfin mentionner à propos de la propédeutique scientifique la capacité de communiquer et d'appliquer des savoirs théoriques (cf. Huber, 1998, p. 161; KMK, 1978, p. 561; Köller et al., 2004, p. 117; Schmidt, 1994, p. 226 ss).

Huber (1997, p. 348) distingue trois niveaux dans la propédeutique scientifique:

1. l'apprentissage et la pratique du travail scientifique, c'est-à-dire la familiarisation avec ses grandes notions et méthodes;
2. l'apprentissage de la science, c'est-à-dire d'une attitude fondée sur le questionnement constant, la démonstration et l'appréciation;
3. l'apprentissage sur la science, c'est-à-dire la réflexion critique dans un contexte plus large.

Huber (1998, p. 150) indique qu'une propédeutique scientifique est impossible sans spécialisation ni travail approfondi sur une question ou un problème. Dans cette optique, le centrage scientifique de l'enseignement peut être compris comme formant la base de traitement de savoirs experts (Tenorth, 1994, p. 176). La propédeutique scientifique englobe par ailleurs les premiers efforts d'acquisition de savoirs experts.

Les objectifs formulés dans le RRM montrent clairement que la propédeutique scientifique ne saurait être exclusivement considérée comme une préparation aux études supérieures, et qu'elle contribue – dans le contexte du gymnase – à la formation de la personnalité (cf. Schmidt, 1994, p. 223), ce qui fait apparaître un recoupement avec les compétences transversales.

1.2.3 Compétences transversales

La question de savoir ce que sont en fait les compétences transversales n'a toujours pas de réponse définitive (Notter & Arnold, 2003, p. 12); on trouve dans la littérature spécialisée divers catalogues et listes à ce sujet. L'étude *Objectifs pédagogiques transversaux* (EVAMAR I, sous-projet 2), par exemple, distingue quatre champs de compétences transversales (Grin et al., 2004, p. 40):

1. les compétences de critique culturelle recouvrent la capacité à discerner des points de vue idéologiques, à reconnaître ses propres préjugés, à se montrer ouvert à d'autres cultures et à envisager sa propre culture sous d'autres angles, ainsi qu'à apprécier des formes artistiques différentes;
2. les compétences de traitement de l'information consistent à savoir où et avec quels outils trouver l'information, à recourir convenablement à divers médias pour se procurer l'information, et à se montrer critique à l'égard de l'application des technologies nouvelles;
3. les compétences scientifiques correspondent à la capacité de mener une argumentation scientifique, de formuler des théories, de concevoir des hypothèses et de les vérifier, mais aussi d'adapter son expression au contexte et de s'exprimer avec clarté et précision;
4. les compétences de coopération permettent à la personne de se prendre en charge et d'assumer ses responsabilités à l'égard des autres, et de travailler individuellement ou en groupe.

Bieri Buschor et Forrer (2005, p. 42) définissent d'une façon tout à fait générale les compétences transversales comme la capacité d'une personne à agir convenablement et avec conscience de ses responsabilités dans diverses situations et dans ses divers rôles sociaux. L'apprentissage transversal ne porte donc plus sur des compétences spécifiquement acquises dans telle ou telle matière. Ce qui ne veut pas dire que les compétences transversales remplacent l'apprentissage spécialisé, bien au contraire: «la recherche montre clairement que l'émergence de compétences pluridisciplinaires implique l'acquisition préalable de solides connaissances spécialisées» (Klieme et al., 2003, p. 75). EVAMAR I a montré que les enseignants qui soutiennent les compétences transversales pensent

qu'elles nourrissent le développement personnel de leurs élèves et les équipe pour des études supérieures ou professionnelles ultérieures (Ramseier et al., 2005, p. 13).

Notter & Arnold (2003, p. 12) ont interrogé des étudiants sur leurs compétences transversales au moment de leur maturité. Ce qui a fait ressortir qu'ils se sentaient le mieux formés à la compréhension de textes (compréhension et interprétation), au travail autonome, à l'expression écrite et à l'apprentissage autonome au moment de leur maturité. Ils se sentaient en revanche le moins bien préparés sur l'analyse éthique des questions, l'exposé en public et la gestion du temps (Notter & Arnold, 2003, p. 23). Il est aussi apparu que les étudiants de toutes les disciplines considéraient comme relativement importantes toutes les compétences transversales, en nommant particulièrement les suivantes:

- apprendre de manière autonome;
- travailler de manière autonome;
- assumer sa responsabilité de son propre apprentissage;
- gérer son temps;
- gérer son stress;
- gérer les problèmes: les identifier, rechercher et élaborer des solutions (Notter & Arnold, 2003, p. 39).

Bieri Buschor et Forrer (2005) ont analysé les compétences transversales de jeunes adultes en Suisse en s'appuyant sur des données d'enquêtes. Leur étude s'efforce de montrer si ces jeunes gens se sentent convenablement équipés pour aborder dans de bonnes conditions les exigences générales de la post-adolescence, dans leur vie publique comme privée. Les catégories de compétences retenues allaient des compétences personnelles à la responsabilité sociale et environnementale et à la compétence politique, en passant par les capacités de coopération et d'apprentissage. On s'aperçoit très vite que cette classification produit des recoupements importants: il est par exemple difficile de distinguer clairement dans l'analyse la réflexion personnelle et les stratégies d'apprentissage cognitif.

1.2.4 Bilan partiel

L'introduction du travail de maturité était sous-tendue par la volonté que les élèves des gymnases soient à même de faire face aux nouvelles exigences actuelles de souplesse et d'organisation personnelle. Le travail de maturité doit donc leur servir à acquérir ces compétences, mais aussi à administrer la preuve qu'ils les possèdent. En ce qui concerne la délimitation de la fonction pédagogique du travail de maturité, il apparaît que jusqu'à présent, les notions de maturité nécessaire aux études supérieures et de propédeutique scientifique n'ont pas été définies de façon uniforme, ni non plus les compétences transversales. On observe un vague consensus sur le fait que les élèves, au cours de la préparation de leur travail de maturité, travaillent individuellement ou en groupe pendant une période relativement prolongée sur un projet, qu'ils approfondissent certaines méthodes de recherche, d'analyse et d'argumentation, et se montrent capables de traiter convenablement des informations scientifiques élémentaires. Le travail de maturité ne doit pas être seulement un effort de production de savoir centré sur la réception, mais aussi une «performance» (au sens linguistique du terme): le savoir acquis au cours du processus doit être communiqué de façon appropriée, sous forme de rapport écrit et de présentation orale. Ce qui veut dire que cet élément important d'émergence de la maturité et des aptitudes requises pour les études supérieures ne remplit pas seulement une fonction propédeutique scientifique au sens étroit, mais qu'il doit contribuer à enrichir les compétences transversales de l'élève, dont le sens de ses responsabilités et la maîtrise des tâches.

1.3 Etat de la recherche sur le travail de maturité

Un petit nombre d'études ont déjà été consacrées à certains aspects du travail de maturité. Dans le cadre du sous-projet 2 de la première phase d'EVAMAR, quelque 3 500 enseignants et 13 000 élèves en dernière année de maturité ont été consultés à l'aide d'un questionnaire standardisé sur les objectifs transversaux (Pagnossin et

al., 2005). Il existe aussi une présentation comparative des cadres légaux et réglementaires du travail de maturité gymnasiale et du travail interdisciplinaire centré sur un projet de la maturité professionnelle (Zillig, 2004). On trouve par ailleurs pour les cantons d'Argovie (Binder & Feller-Länzlinger, 2004), de Genève (Davaud & Hexel, 2003), de Lucerne (Binder & Feller-Länzlinger, 2003) et de Vaud (Stocker, 2006) des rapports d'évaluation de la mise en œuvre du RRM 95 comportant des indications sur le travail de maturité. Et on dispose aussi pour un gymnase fribourgeois de deux articles sur des enquêtes internes relatives au travail de maturité (Darbellay, 2002; Schmid, 2002). Nous résumons dans les pages qui suivent les principaux résultats de ces travaux.

1.3.1 Typologie des travaux de maturité

Le RRM 95 ne donne aucune précision sur le contenu ni la forme du travail de maturité. L'étude de Pagnossin et al. (2005) fournit toutefois quelques indications sur ce qui se fait dans la pratique. On distingue ainsi cinq grands types de travaux de maturité:

- travaux théoriques (41%);
- travaux empiriques (33%);
- productions artistiques, audiovisuelles et concrètes (12%);
- activités extrascolaires (12%);
- projets socioculturels (2%).

Cette classification n'est pas très nette: on peut par exemple arguer qu'une production artistique présente une composante empirique.

La répartition des types de production révèle de claires différences entre les régions linguistiques. Les travaux empiriques sont les plus demandés en Suisse alémanique, alors que la Suisse italienne et la Suisse francophone leur préfèrent les travaux théoriques. L'art et l'audiovisuel sont sensiblement plus fréquents en Suisse alémanique qu'en Suisse latine. En ce qui concerne le sexe, on n'observe que deux différences notables: les jeunes gens s'orientent plus volontiers que les jeunes filles vers les travaux théoriques (42,5% contre 39%) et ces dernières choisissent plus souvent des projets socioculturels ou extrascolaires (11,5% contre 16%). Les auteurs de ce sous-projet EVAMAR imputent les variations entre régions linguistiques aux différences entre les dispositions cantonales. Les écarts entre les sexes ne font pas ici l'objet d'une étude plus poussée, compte tenu de leur faible ampleur (Pagnossin et al., 2005, p. 185 s).

1.3.2 Choix du sujet

Zillig (2004, p. 15 s) constatait dans son étude que les pratiques varient d'une école à l'autre en ce qui concerne le choix du sujet. Les élèves ont dans l'ensemble en Suisse alémanique plus de liberté qu'en Suisse francophone et italophone. Ce qu'a confirmé l'étude de Pagnossin et al. (2005): 97% des élèves alémaniques interrogés choisissaient librement leur sujet; en Suisse romande, 65% voyaient leur liberté circonscrite par un thème général ou une liste de sujets. En Suisse italienne, 8% seulement d'entre eux pouvaient librement choisir leur sujet. Ce qui s'explique par le fait qu'au Tessin, le suivi des travaux de maturité est assuré sous forme de séminaires thématiques (Pagnossin et al., 2005, p. 184 s).

La grande majorité des élèves choisissent un sujet qui les intéresse. Mais d'autres facteurs peuvent aussi jouer, comme les connaissances préalables, des tiers ou l'envie de relever un défi (Binder & Feller-Länzlinger, 2003, p. 9; 2004, p. 41; Pagnossin et al., 2005, p. 184 s). L'enquête menée dans un gymnase fribourgeois montre que 61% des élèves interrogés étaient d'accord avec leur sujet (Schmid, 2002, p. 11).

Parmi les élèves interrogés dans l'enquête EVAMAR I, 24% seulement avaient pris un sujet couvrant plus d'une discipline, 39% un sujet portant sur une seule matière, et 38% un sujet dans une discipline extrascolaire (Pagnossin et al., 2005, p. 186). Il faut remarquer à ce propos que le cadre en place dans chaque école peut jouer

à ce niveau un rôle considérable, d'autant que les établissements n'admettent pas tous les sujets sortant des matières enseignées (Zillig, 2004, p. 15 s).

1.3.3 Réalisation

La grande majorité des élèves préparent leur travail de maturité seuls (Pagnossin et al., 2005, p. 191; Binder & Feller-Länzlinger, 2003). Font exception ceux du canton d'Argovie, où 76% d'entre eux réalisent leur travail de maturité en équipe. Binder et Feller-Länzlinger expliquent cet écart par les prescriptions du canton d'Argovie, qui soulignent que le travail de maturité doit être préparé en groupe, mais également par le fait que le travail préalable centré sur un projet est généralement aussi réalisé en équipe (Binder & Feller-Länzlinger, 2004, p. 42 s). Les auteurs se sont enquis dans les cantons de Lucerne et d'Argovie du temps consacré par les élèves à leur travail de maturité. Ils observent qu'il varie énormément d'un gymnase à l'autre, et qu'il a tendance à être plus long dans le canton de Lucerne que dans celui d'Argovie; cela pourrait être dû à leur avis au fait que les élèves argoviens travaillant en équipe, ils ont ensuite besoin de moins de travail individuel.

Dans la plupart des établissements, le travail de maturité prend de 10 à 18 mois (Zillig, 2004, p. 17). Interrogés sur le temps qui devrait leur être imparti entre le démarrage (contrat de travail ou de projet) et la remise du travail de maturité, la plupart des élèves de Lucerne et d'Argovie ont répondu six mois (Binder & Feller-Länzlinger, 2003, p. 12; 2004, p. 44). Une période de un an a été jugée excessive dans l'enquête menée au sein d'un gymnase fribourgeois; une nette majorité (63%) des enseignants et des élèves interrogés dans cet établissement ont toutefois estimé qu'il suffisait de moins de huit mois (Schmid, 2002, p. 10).

Diverses écoles prévoient des plages d'enseignement pour le travail de maturité, par exemple des leçons intégrées dans l'horaire, des semaines bloquées, etc. (Zillig, 2004, p. 18). Interrogés sur le temps à prévoir pour cela, 28% des élèves lucernois ont souhaité de un à cinq jours de cours, 32% de six à dix, et 19% de onze à quinze (Binder & Feller-Länzlinger, 2003, p. 14).

1.3.4 Savoirs, compétences et savoir-faire des élèves

Deux tiers des enseignants interrogés dans l'étude EVAMAR I estimaient que les élèves étaient suffisamment capables de mener à bien un travail de manière autonome (Pagnossin et al., 2005, p. 177). Et 91% des élèves eux-mêmes partageaient cet avis (Pagnossin et al., 2005, p. 188). En ce qui concerne le bagage méthodologique des élèves, la moitié seulement des enseignants interrogés jugeaient qu'il suffisait à la préparation autonome d'un travail de maturité, contre sept élèves sur dix de cet avis (Pagnossin et al., 2005, pp 177 et 188).

Les élèves consultés dans le gymnase de Fribourg se sont déclarés moins satisfaits; la moitié environ de la section francophone de l'établissement et 58% de la section germanophone ont jugé avoir été insuffisamment préparés au travail scientifique (Darbellay, 2002, p. 14; Schmid, 2002, p. 11).

L'évaluation de l'application du RRM 95 dans le canton d'Argovie a révélé que les élèves se montrent plutôt critiques à l'égard du travail centré sur un projet effectué en préparation au travail de maturité: à leurs yeux, les savoir-faire transmis à cette occasion n'auraient qu'une utilité limitée dans la préparation du travail de maturité (Binder & Feller-Länzlinger, 2004, p. 45).

1.3.5 Suivi du travail de maturité

La préparation du travail de maturité fait l'objet d'un suivi dans tous les établissements, mais les pratiques varient en la matière (Zillig, 2004, p. 17). La majorité (80%) des élèves lucernois ont eu de une à huit réunions avec l'enseignant chargé de les encadrer; ces entretiens ont surtout porté sur des conseils sur le fond, des questions de méthodologie scientifique, le choix du sujet, la présentation du travail et la rédaction du rapport (Binder & Feller-Länzlinger, 2003, p. 17 s). Dans l'enquête interne menée dans le gymnase germanophone fribourgeois, 47% des élèves interrogés ont estimé que le suivi du travail de maturité devait davantage revêtir la forme d'entretiens; il est apparu en outre qu'un tiers environ des élèves se sont sentis limités par l'encadrement, mais que 13% d'entre eux auraient aimé recevoir plus de consignes (Schmid, 2002, p. 10).

Dans l'enquête nationale de Pagnossin et al. (2005, p. 177 ss), 54% des enseignants interrogés se sont déclarés satisfaits du temps dont ils disposent pour encadrer les travaux de maturité, et 43% ont jugé suffisantes les décharges horaires qui leur sont attribuées. Il ressort de l'expérience réunie dans le gymnase germanophone de Fribourg qu'une heure de décharge ne suffit pas pour un séminaire de huit élèves: pour que le solde soit nul entre le travail et ses bénéfices, il faudrait une décharge d'une heure pour un maximum de cinq ou six élèves (Schmid, 2002, p. 11).

Les élèves brossent un tableau différent – du moins dans le canton de Vaud, où trois quarts de ceux qui ont été interrogés sont satisfaits ou très satisfaits du temps consacré à leur encadrement (Stocker, 2006, p. 35). De même, la majorité des élèves interrogés dans EVAMAR I jugeaient l'encadrement suffisant, alors que les enseignants se plaignaient de la pénurie de temps (Pagnossin et al., 2005, p. 177 s). L'étude lucernoise confirme ce résultat chez les élèves (Binder & Feller-Länzlinger, 2003, p. 16).

La plupart des élèves comme des enseignants jugent suffisantes les ressources matérielles dont ils disposent. On remarque toutefois dans l'étude de Pagnossin et al. que les opinions divergent entre les enseignants et les élèves en ce qui concerne les objectifs et les conditions de réalisation du travail de maturité: la plupart des enseignants les trouvent suffisamment bien définis, alors que 55,6% des élèves interrogés ne trouvent pas suffisantes les consignes et directives qui leur sont données (Pagnossin et al., 2005, p. 177 ss). Les résultats sont similaires chez les élèves des gymnases des cantons de Vaud et de Genève: 45% des élèves vaudois interrogés ont trouvé que les directives et consignes n'étaient pas suffisamment claires (cf. Davaud & Hexel, 2003, p. 63; Stocker, 2006, p. 35). Et dans le canton de Genève, Davaud et Hexel constatent aussi que c'est dans les gymnases où les élèves sont plutôt mécontents des conditions de réalisation du travail de maturité que l'utilité de ce dernier a tendance à être mise en doute (Davaud & Hexel, 2003, p. 62 s).

1.3.6 Bilan partiel

Ces études livrent des résultats intéressants sur le travail de maturité. Mais de nombreuses questions restent en suspens: les résultats de l'étude de Pagnossin et al (2005) reposent uniquement sur l'appréciation des enseignants et des élèves interrogés; les travaux eux-mêmes n'ayant pas été examinés, on peut tout au plus en tirer des conclusions qualitatives indirectes. Les travaux de Binder & Feller-Länzlinger (2003, 2004), de Davaud & Hexel (2003) et de Stocker (2006) collectent aussi des données sur le travail de maturité, mais ce sont des rapports d'évaluation d'enquêtes cantonales dont on ne saurait inférer d'une manière générale de conclusions nationales; et l'enquête fribourgeoise se limite même à un unique établissement (Darbellay, 2002; Schmid, 2002). Le travail de Zillig (2004) montre que les conditions de réalisation varient d'un canton et d'un établissement à l'autre, mais il n'examine pas si et dans quelle mesure ces variations ont un impact sur la qualité des travaux de maturité.

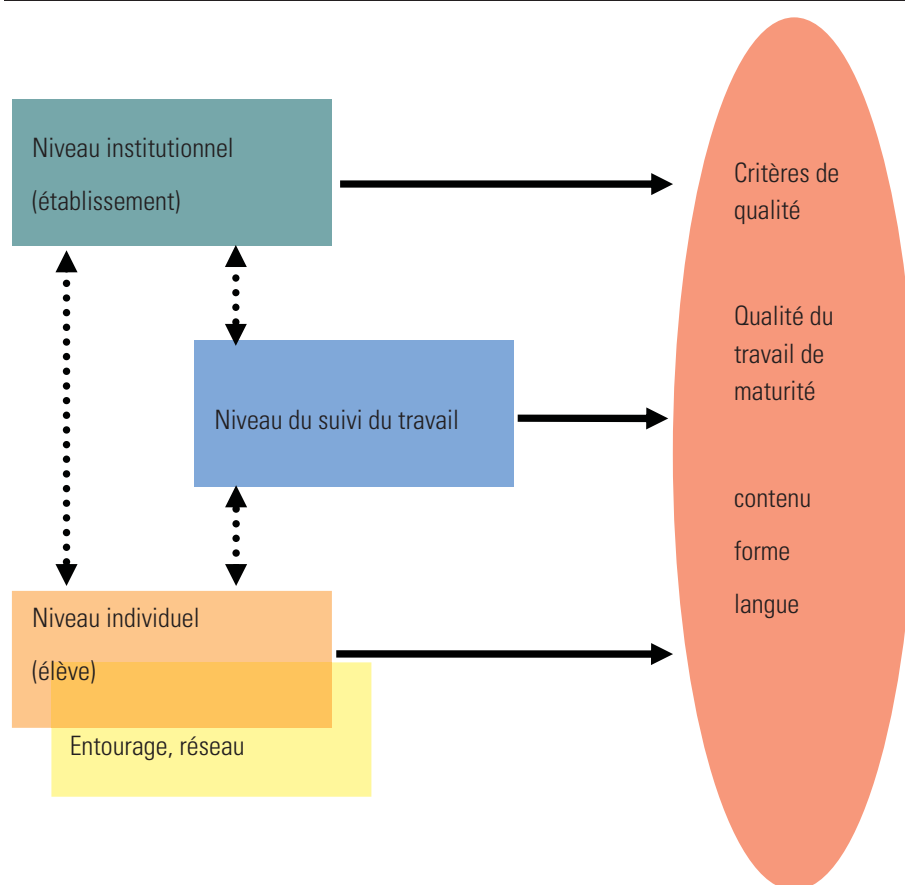
2 Bases méthodologiques et réalisation de l'enquête

Nous allons maintenant présenter la conception de l'enquête d'analyse de la qualité des travaux de maturité. Nous partons d'un schéma de fonctionnement du processus de production du travail de maturité. Nous montrerons comment ont été conçus les instruments de collecte des données, et comment ces dernières ont été concrètement analysées. Nous aborderons aussi le mode d'échantillonnage et la population examinée.

2.1 Schéma de production d'un travail de maturité

La planification de l'enquête et la définition des hypothèses se sont appuyées sur un schéma de fonctionnement (figure VI.1) visualisant les rapports de dépendance à prendre en compte dans l'analyse des travaux de maturité.

Figure VI.1: Genèse d'un travail de maturité



Sur la base de ce schéma, l'enquête a été subdivisée en trois volets. Tout d'abord, l'analyse d'un échantillon national de travaux de maturité selon une grille de critères. Elle a été complétée par l'examen du cadre institutionnel dans lequel les élèves préparent leurs travaux de maturité, cela dans une perspective comparative. Un questionnaire standardisé a d'autre part été envoyé à tous les élèves des classes retenues; il s'agissait de recueillir des données sociodémographiques et des renseignements d'ordre général sur le travail de maturité (le sujet et les modalités de son choix, la note donnée par l'enseignant), mais aussi et surtout sur l'encadrement du travail (forme et fréquence du suivi, satisfaction éprouvée à son égard), et sur l'appui et les ressources dont avait bénéficié l'élève (satisfaction éprouvée à l'égard du cadre scolaire, assistance offerte, réseau social). Les élèves ont par ailleurs été invités à apprécier le volume de travail personnel qu'ils avaient consacré à leur travail de maturité, et l'utilité de ce dernier.

2.2 Evaluation des travaux de maturité

2.2.1 Dimensionnement de l'enquête

L'étude est centrée sur le travail de maturité, produit d'un travail que les données disponibles ne permettent pas de reconstituer dans le détail. Ce qui veut dire qu'il demeure impossible d'intégrer les aspects psychométriques et sociométriques dans l'évaluation de la qualité des travaux de maturité. On ne connaît pas non plus le volume de travail effectivement fourni par les élèves, car il est impossible de quantifier l'aide qu'ils ont par exemple pu recevoir de leur famille ou de leurs connaissances. C'est ce que veut dire Oswald Inglin, sous-directeur du gymnase Leonhard de Bâle, quand il appelle le travail de maturité une joint-venture familiale: «toute une armée de parents, de frères et sœurs et d'amis passent un temps considérable à relire et à améliorer les travaux de maturité. Selon leur situation sociale, certains élèves sont ainsi favorisés, et d'autres défavorisés par leur milieu» (Inglin, 2007, p. 27).

L'évaluation de la qualité d'un travail de maturité ne permet pas de conclure avec certitude à la compétence de son auteur. Il s'agit souvent d'un processus unique, dont le caractère durable ne ressort pas directement du produit: il n'est pas sûr que les compétences qu'il reflète soient déjà bien solides et transférables avec succès dans un autre contexte (comme les études supérieures) ou dans des conditions légèrement différentes, comme un encadrement moins attentif (cf. Bieri Buschor & Forrer, 2005, p. 44). Mais ce centrage sur le travail de maturité vu comme le produit d'un processus présente toutefois l'avantage décisif de prévenir tout effet de halo, la personne chargée de l'évaluation ne sachant rien de l'élève concerné.

2.2.2 Normes de référence servant à juger de la qualité des travaux de maturité

La qualité d'un travail de maturité ne peut s'apprécier qu'en référence à une norme qui serve d'étalon. Dans le cas qui nous occupe, il paraît le plus approprié de s'appuyer sur des normes portant sur le contenu. Une norme individuelle n'est pas envisageable, s'agissant d'évaluer le produit d'un processus que l'on ne peut pas observer. Les normes sociales ont leur place dans cette étude, mais elles laissent des zones d'ombre que l'on ne saurait pourtant négliger. Il n'est tout d'abord possible de procéder à des comparaisons fondées sur des caractéristiques sociales (région linguistique, sexe, etc.) qu'au sein de notre échantillon; mais ce dernier introduisant des distorsions, surtout en ce qui concerne la Suisse romande (voir 2.6 ci-dessous), on court le risque d'arriver à des résultats erronés (cf. Rheinberg, 2002, p. 64 ss). D'autre part, une comparaison au sein de notre échantillon présuppose que l'on a évalué la qualité des travaux de maturité, c'est-à-dire qu'il existe déjà une norme objective d'appréciation du contenu.

On ne possède pour l'instant pas d'instruments permettant d'évaluer les travaux de maturité en référence à des normes objectives. On trouve toutefois dans la littérature des modèles et des catalogues de compétences acquises ou enrichies à l'occasion de la préparation d'un travail écrit. Mais ils sont souvent peu systématiques, voire arbitraires (cf. par exemple Bonati & Hadorn, 2007; Wilhelmer, 2005, p. 85; Winter, 2006), et se prêtent donc mal à notre recherche.

Les diverses exigences auxquelles doivent répondre les élèves préparant leur travail de maturité donnent déjà des indications sur ces normes objectives de contenu, par exemple:

1. les exigences des établissements scolaires (objectifs de la formation gymnasiale);
2. les exigences des établissements accueillant ensuite les bacheliers dans l'enseignement supérieur et la formation professionnelle (aptitude aux études supérieures, qualifications clés);
3. les exigences liées à chaque discipline (caractère scientifique);
4. les exigences langagières.

1) Exigences des établissements scolaires

Les objectifs de formation des établissements de formation gymnasiale sont définis dans le RRM 95 (CDIP, 1995) et le plan d'études cadre des écoles de maturité (CDIP, 1994). Aucun de ces documents ne contient de mention concrète du travail de maturité ni des objectifs que doit permettre d'atteindre sa préparation. Les objectifs qu'ils définissent donnent toutefois des pistes générales de formulation de normes relatives au contenu du travail (cf. sections 1.2 ci-dessus et 3.1.1 ci-dessous).

L'étude EVAMAR I examinait les aspects qualitatifs et les critères d'évaluation auxquels les enseignants et les élèves des gymnases accordent une importance particulière. Les critères d'évaluation avaient été regroupés en cinq catégories:

1. démarche scientifique;
2. apport individuel de l'élève;
3. présentation du travail;

4. caractère interdisciplinaire du travail;
5. originalité du travail.

Presque tous les aspects avaient été jugés importants. Plus de 80% des enseignants interrogés avaient ainsi estimé importants la démarche scientifique, l'apport individuel de l'élève et la présentation. Le caractère interdisciplinaire du travail et son originalité avaient été toutefois moins souvent retenus comme importants (< 50% pour le caractère interdisciplinaire, et 53,5% pour l'originalité). Parmi les élèves, ce sont le travail personnel et la présentation qui avaient été en général considérés comme importants, mais moins souvent la démarche scientifique (cf. Pagnossin et al., 2005, p. 179 s).

La présente étude ne s'appuyant que sur le produit final, c'est-à-dire le travail de diplôme sous forme de document textuel, elle ne peut pas prendre en compte l'apport personnel de l'élève ni la présentation orale, pourtant qualifiés d'importants. En revanche, la démarche scientifique (familiarisation avec la méthodologie scientifique), qui figure aussi à l'article 5 du RRM 95, a été prise comme norme de référence dans notre étude (cf. section 2.2.3).

Les exigences des établissements de formation gymnasiale ont par ailleurs été prises en considération sous forme de synoptique des grilles d'évaluation utilisées dans un certain nombre d'écoles prises au hasard. C'est de ce synoptique que nous sommes en fin de compte partis pour concevoir notre propre instrument d'évaluation.

2) Exigences des établissements accueillant ensuite les bacheliers

Aux yeux des établissements accueillant ensuite les bacheliers, ce sont surtout les qualifications clés et les compétences transversales qui comptent (voir également section 1.2). Or, nous l'avons dit, il est impossible de conclure à l'acquisition effective de compétences sur la base du produit d'un travail unique; c'est pourquoi l'analyse de cette dimension transversale n'est pas possible non plus. La préparation et la présentation d'un travail de maturité sont nécessaires, mais ne montrent pas à elles seules qu'un élève est apte aux études supérieures – d'autant plus que les universités elles-mêmes ne s'appuient pas sur une définition universelle de la teneur de l'aptitude aux études supérieures. Cela dit, les établissements universitaires ont tous les mêmes exigences en matière d'expression écrite, de traitement des données et des faits, de respect des règles méthodologiques, de pensée analytique indépendante et de logique de l'argumentation, qui se reflètent dans les directives de préparation des travaux écrits à l'université (cf. à titre d'exemples Erdfelder, Musch & Cüppler, 2002; Carolus, 2006; Wallmeier, 2007; ainsi que Konegen-Grier, 2002).

3) Exigences propres à chaque discipline

Au premier point (établissements produisant les bacheliers), on a vu que les enseignants considèrent la démarche scientifique comme particulièrement importante. Au second (établissements accueillant les bacheliers), le caractère scientifique est apparu comme un trait institutionnel primordial. Mais chaque discipline possède son approche de la théorisation et de la méthodologie, et s'inscrit dans une tradition différente de professionnalisation. Il est donc impératif de faire juger chaque travail de maturité par un spécialiste au moins, ce qui garantira que le caractère scientifique aura été apprécié selon les normes en vigueur dans la discipline.

4) Exigences langagières

Il existe divers instruments d'analyse langagière des textes, comme la grille d'analyse textuelle de Zurich (*Zürcher Textanalyseraster*, Nussbaumer & Sieber, 1994), mais ils se concentrent avant tout sur les aspects linguistiques du texte, et ne se prêtent donc pas très bien à l'évaluation des travaux de maturité. La grille de Zurich permettant de relever les qualités et les défauts des textes, et donc de comprendre leur cohérence dans une approche (re)constructive (Nussbaumer & Sieber, 1994, p. 185), elle a toutefois fourni de précieux apports à la mise au point de la grille d'analyse des travaux de maturité.

La grille de Zurich comporte trois volets.

- Volet 0: valeurs de référence et corrélats

Ce volet mesure surtout des caractéristiques d'un texte, comme sa longueur, la richesse et la variété du vocabulaire et de la syntaxe, l'ampleur et la nature de la panoplie de procédés d'expression de la cohésion, la difficulté du sujet abordé dans le texte (Nussbaumer & Sieber, 1994, p. 156 s).

- Volet A: respect du système de la langue et orthographe

Ce volet examine le respect des règles d'orthographe, de ponctuation, de grammaire (morphologie, syntaxe et structure du texte) et de sémantique; il s'agit donc de la correction formelle du texte (Nussbaumer & Sieber, 1994, p. 160 s)

- Volet B: bon usage des outils langagiers

Il ne s'agit plus ici, comme dans le volet A, d'un jugement en juste ou faux, mais d'une appréciation selon une échelle allant d'un emploi particulièrement heureux à un emploi particulièrement malheureux (Nussbaumer & Sieber, 1994, p. 161 ss).

Pour l'analyse des travaux de maturité, on a délibérément exclu l'examen des caractéristiques du texte prévu dans le volet 0 de la grille de Zurich, l'objectif étant de porter une appréciation qualitative sur les travaux de maturité examinés, et non de mesurer des aspects quantitatifs relevant de l'expression écrite.

Les questions de correction (orthographe, ponctuation, morphologie, syntaxe, structure et sémantique), et le bon usage des outils langagiers (volets A et B de la grille de Zurich) ont toutefois une importance dans la détermination de la qualité des travaux de maturité. L'étude relevant plus des sciences sociales que de la linguistique, les aspects formels de l'utilisation de la langue n'ont été jugés que sur des critères généraux (continuité du texte, clarté de l'expression, choix et variété du vocabulaire, expressivité, structure syntaxique) et de correction.

Nussbaumer et Sieber (1994, p. 167) soulignent que certains critères suscitent des difficultés particulières d'objectivité, et que l'évaluation d'un même texte sur ces critères (ils mentionnent par exemple la qualité des ressources langagières utilisées) peut diverger considérablement d'un correcteur à l'autre. Ces difficultés se retrouvent aussi dans l'étude de la qualité des travaux de maturité. Il est extrêmement difficile de mesurer la fiabilité des évaluateurs entre eux du fait qu'ils ont opéré en binômes changeants, et pour des parts variables; pour la consistance de leur jugement, nous avons donc pris comme critère de contrôle la corrélation intraclasse des appréciations globales.

La constatation pragmatique que chaque type de texte doit répondre à des exigences fonctionnelles différentes a aussi débouché sur d'autres indices. Kruse (1999) distingue quatre exigences auxquelles doit répondre un texte.

1. Précision: c'est la précision avec laquelle le texte exprime l'intention de son auteur.
2. Fonction de communication: elle ressort de l'impact social que doit avoir un texte.
3. Structuration: c'est la structure interne du texte, avec les liens entre ses divers éléments.
4. Authenticité affective: il s'agit de la mesure dans laquelle l'expressivité d'un texte reproduit bien le vécu affectif de la personne qui l'a produit.

Le quatrième aspect, celui de l'authenticité affective, n'est pas très important en ce qui concerne les travaux de maturité. Les trois autres, en revanche, le sont tout à fait. La grille de critères a donc été articulée autour des trois types d'exigences (fonction de communication, structuration et précision), définis par Kruse.

2.2.3 Conception de l'instrument d'analyse

Sur la base de cette réflexion sur les aspects des travaux de maturité à évaluer, a été conçue une grille de critères selon un processus circulaire, par induction comme déduction; elle a ensuite été testée à plusieurs reprises, jusqu'à ce qu'elle paraisse appropriée et utilisable. Ces nombreux contrôles de fonctionnalité et de validité ont bien sûr permis d'optimiser la grille, mais ils ont aussi familiarisé parfaitement ses concepteurs avec son maniement.

La grille a ensuite été opérationnalisée sur une échelle à six degrés. L'intention initiale était d'adopter une échelle à quatre degrés, pour opérationnaliser chaque point; mais cela s'est révélé difficile, d'autant qu'il n'a pas été possible de définir et de distinguer chaque niveau avec suffisamment de précision et de clarté. Outre la difficulté de décrire verbalement chaque niveau, les tests ont révélé que la subdivision en quatre degrés était trop grossière pour refléter les écarts de qualité entre les travaux. Il avait été décidé de ne pas offrir de réponse médiane neutre, et c'est pourquoi on a finalement opté pour une échelle à six degrés, avec verbalisation des deux items extrêmes.

2.2.4 Notation

Chaque travail de maturité a été examiné par deux évaluateurs, dont l'un au moins possédait toujours des compétences spécifiques dans la discipline dont relevait le travail de maturité. Pour les travaux interdisciplinaires, on a veillé dans la mesure du possible à couvrir toutes les spécialités concernées. Les experts étaient en général titulaires d'un diplôme d'études supérieures dans l'une ou plusieurs des disciplines touchées par le travail de maturité; ces personnes possédaient aussi pour la plupart une solide expérience de la pédagogie de la discipline dans le secondaire II. Pour la seconde notation, on a eu recours à des personnes qui possédaient des qualifications en sciences de l'éducation et en sciences sociales, plusieurs années d'expérience dans la recherche, et en général aussi des compétences dans l'enseignement universitaire.

L'expérience réunie dans la mise au point de la grille d'analyse a montré qu'il était indispensable de former les évaluateurs si l'on voulait obtenir la meilleure uniformité possible de notation entre eux, et ainsi une bonne objectivité. Au cours de cette formation, ils ont tous reçu un guide d'utilisation de la grille de critères.

Pour l'évaluation, les travaux avaient été rendus anonymes; il était donc impossible que les évaluateurs prennent en compte d'autres facteurs, comme le sexe de l'élève.

2.2.5 Contrôle anti-plagiat

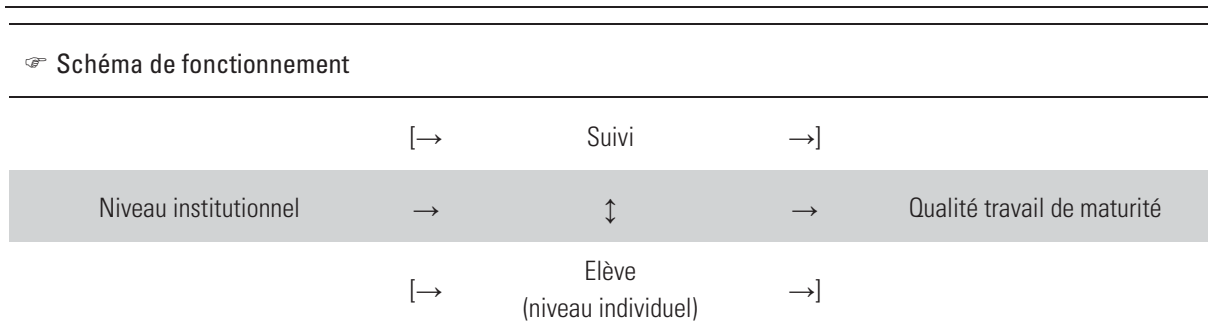
Il a été demandé aux experts d'apprécier si les travaux qui leur étaient soumis risquaient de contenir des plagiat. Un doute émis par l'un d'eux ne signifiait bien sûr pas nécessairement que le travail était en faute, et l'absence de doute exprimé ne signifiait pas non plus automatiquement qu'il ne l'était pas. C'est pourquoi deux fois dix travaux ont été choisis au hasard parmi ceux qui avaient et n'avaient pas éveillé de soupçons, et ont été soigneusement contrôlés.

Sur ces vingt travaux soumis à un contrôle anti-plagiat approfondi, treize avaient été remis sous forme électronique, et il a donc été possible d'utiliser le logiciel *Plagiarism Finder*⁴⁷. En fin de contrôle, *Plagiarism Finder* présente une statistique donnant notamment le pourcentage de correspondances de mots trouvés sur l'Internet. Les endroits douteux sont directement repérés dans le document, avec le lien hypertexte, de sorte qu'il est possible de consulter les textes de référence pour déterminer s'il s'agit ou non d'un plagiat. Dans un second temps, ces vingt travaux ont aussi été soumis à la «technique Google» que recommande Weber Wulff (sans date), c'est-à-dire qu'une recherche a été faite dans Google sur des passages suspects pour déterminer s'ils auraient été recopiés de l'Internet.

⁴⁷ *Plagiarism Finder* analyse le texte et le segmente en mots. Après quoi, il réunit les mots en séquences (par exemple de sept mots consécutifs), qu'il vérifie sur un moteur de recherche sur l'Internet.

2.3 Comparaison des conditions de réalisation des travaux

Figure VI.2: Les conditions de réalisation des travaux de maturité dans le schéma de fonctionnement



Il ressort du schéma de la figure 1 que le niveau institutionnel peut avoir une influence directe comme indirecte sur la qualité des travaux de maturité. Il convient dans ce contexte de distinguer les directives nationales, cantonales, et propres à l'établissement.

A l'échelon national, l'analyse a porté sur le RRM 95 et le plan d'études cadre. Les gymnases sont tenus de respecter le RRM 95, mais pas nécessairement le plan d'études cadre, qui a simplement valeur de recommandation formelle à l'intention des cantons en ce qui concerne la conception de leurs plans d'études gymnasiales.

Les dispositions cantonales relatives aux travaux de maturité figurent à divers niveaux de la législation. C'est pourquoi l'analyse des conditions de réalisation a porté sur toutes les lois et ordonnances cantonales, et a visé à déterminer si elles contenaient des indications relatives au travail de maturité. La même recherche a été faite dans les plans d'études cantonaux.

La présentation des conditions de réalisation définies par les établissements eux-mêmes se fonde sur des analyses de contenu des informations figurant dans les documents écrits (directives, règlements, guides, consignes, etc.) émanant des établissements et en vigueur pour l'année 2006-2007, à laquelle se rapportait l'échantillon. Les documents transmis par les écoles ont été anonymisés.

2.4 Enquête auprès des élèves

Figure VI.3: Les élèves dans le schéma de fonctionnement



Le questionnaire standardisé administré aux élèves pour le troisième volet de l'enquête couvrait les catégories d'informations suivantes:

- variables sociodémographiques;
- indications relatives au travail de maturité (discipline, mode de choix du sujet, évaluation);

- suivi (forme et fréquence, satisfaction éprouvée à l'égard de l'appui reçu);
- assistance, ressources (préparation et assistance offertes, satisfaction éprouvée à l'égard des conditions de réalisation du travail prévues par l'école);
- évaluation de l'effort investi et de l'utilité du travail de maturité.

Pour la préparation du questionnaire, on s'est reporté à l'enquête EVAMAR I effectuée auprès de classes terminales (cf. Ramseier et al., sans date). La plupart des questions posées à l'époque sur le travail de maturité ont été reprises, parfois sous une forme légèrement différente, de façon à obtenir une perspective longitudinale. Comme l'examen d'un travail de maturité ne permet pas de déterminer s'il a effectivement contribué la réalisation des objectifs de formation définis dans le RRM, il a été en outre demandé aux élèves dans quelle mesure leur travail de maturité les avait aidés à atteindre ces objectifs.

2.5 Vérification de l'hypothèse de fonctionnement

À la fin des trois volets de l'enquête, les données ont été combinées de sorte qu'il soit possible de discerner dans quelle mesure les fonctionnements dont on avait fait l'hypothèse dans le schéma correspondaient bien à la réalité. L'hypothèse sous-jacente couvrait les quatre champs figurant dans le schéma.

1. Qualité du travail de maturité
2. Niveau institutionnel:
 - a) choix du sujet;
 - b) préparation;
 - c) conditions de réalisation.
3. Niveau du suivi
4. Niveau individuel:
 - a) motivation et perception de soi;
 - b) valeur intrinsèque et extrinsèque du travail.

Nos hypothèses relatives à la qualité des travaux de maturité se sont appuyées sur la réflexion présentée ci-dessus à-propos du développement des compétences transversales. Il en est allé de même pour le niveau individuel. En ce qui concerne le niveau institutionnel et le suivi, nous nous sommes fondés sur les renseignements donnés par le gymnase sur ses objectifs de formation.

2.5.1 Qualité du travail de maturité

Dans la présente étude, nous avons fait l'hypothèse qu'il est judicieux de distinguer dans les travaux de maturité la qualité du contenu, celle de la forme et celle de la langue, et qu'il existe entre ces trois champs qualitatifs des liens importants, qui se reflètent dans l'évaluation (→ hypothèse 1a). Nous avons également fait l'hypothèse qu'il existe des écarts liés au sexe, à la région linguistique et au mode de travail (solitaire ou en groupe) dans ces trois champs qualitatifs (→ hypothèse 1b), mais que le fait qu'un travail relève d'une discipline unique ou soit interdisciplinaire n'a pas d'influence (→ hypothèses 1c et d).

Tableau VI.1: Hypothèses sur les aspects généraux de la qualité

1	Aspects généraux de la qualité
1a	Il existe une corrélation positive entre la qualité du contenu, celle de la forme et celle de la langue des travaux de maturité.
1b	Il existe des écarts significatifs liés au sexe, à la région linguistique et à la forme de travail dans la qualité du contenu, de la forme et de la langue.
1c	La discipline dont relève un travail de maturité n'a pas d'influence significative sur les divers aspects de sa qualité.
1d	Les travaux interdisciplinaires ne présentent pas de différence de qualité significative avec les travaux relevant clairement d'une discipline.

2.5.2 Niveau institutionnel

On peut faire l'hypothèse que les conditions dans lesquelles les écoles placent leurs élèves préparant un travail de maturité ont un impact sur la qualité de ce dernier. L'étude de Zillig (2004) avait déjà montré que ces conditions varient d'un établissement à l'autre – par exemple pour ce qui est des modalités imposées de choix du sujet.

Choix du sujet

Dans la présente étude, nous sommes partis de l'idée que la liberté de choix du sujet a un impact sur la qualité du contenu du travail de maturité. Si l'on estime qu'une grande liberté de choix se traduit par un intérêt et une motivation plus intenses chez l'élève, on peut faire l'hypothèse qu'une grande liberté de choix du sujet aura une influence positive sur la qualité du contenu du travail (→ hypothèse 2a).

Tableau VI.2: Hypothèse relative au choix du sujet

2a	Choix du sujet
	Plus l'élève est libre dans le choix de son sujet, meilleure devient la qualité du contenu du travail de maturité.

Préparation

La préparation offerte a aussi sa place dans les conditions réunies au niveau institutionnel. Elle englobe par exemple des semaines de travail sur projet ou de travail intensif sur le travail de maturité, des cours de rédaction scientifique, l'apprentissage par projet ou des travaux de semestres antérieurs. On peut penser que ces offres ont un effet positif sur la qualité du travail de maturité, en particulier pour ce qui est de sa forme (→ hypothèses 2b1 et 2b2). La qualité des préparations de ce type pouvant varier considérablement d'une école ou d'un enseignant à l'autre, il convient d'examiner s'il existe une corrélation positive entre l'utilité perçue de ces formations et la qualité des travaux de maturité (→ hypothèse 2b3).

Tableau VI.3: Hypothèses relatives à la préparation

2b	Préparation
2b1	Une préparation intense (semaine de travail intensif ou sur un projet, travail par projet, initiations, etc.) a un effet positif sur la qualité du travail.
2b2	Les travaux de maturité des élèves qui n'ont assisté qu'à des séances d'information ou n'ont même suivi aucune préparation ont tendance à présenter une qualité inférieure.
2b3	Les séances de préparation ressenties comme utiles ont un effet positif sur la qualité.

Conditions de réalisation

L'étude de Zillig (2004) avait montré que la plage de variation de la période de préparation du travail de maturité atteint une dizaine de mois. La plupart des élèves interrogés par Binder et Feller-Länzlinger (2003; 2004) avaient estimé que la durée idéale de préparation était de six mois. De même, il était ressorti de l'enquête de Schmid (2002) que les élèves trouvaient trop long un délai d'une année, et qu'il suffisait à leurs yeux de moins de huit mois. Ce qui conduit à faire l'hypothèse qu'il serait possible de déterminer un délai idéal de préparation sur la base de la qualité des travaux de maturité (→ hypothèse 2c1).

Diverses études ont révélé que de nombreux élèves ne trouvent pas satisfaisantes les consignes, directives et instructions que leur donne leur école (cf. Davaud & Hexel, 2003; Pagnossin et al., 2005; Stocker, 2006). Davaud et Hexel (2003) avaient en outre constaté que les élèves plutôt mécontents des conditions dont ils ont bénéficié sont aussi ceux qui ont tendance à mettre en doute l'utilité du travail de maturité. Il convient donc d'examiner si la transparence des consignes, et en particulier celle des critères d'évaluation, se traduit par une meilleure qualité des travaux de maturité (→ hypothèses 2c2, 2c3, 2c4). Outre la transparence des consignes et des critères d'évaluation, il faut aussi se demander si la satisfaction éprouvée à cet égard a un impact sur la qualité des travaux de maturité (→ hypothèse 2c5).

Tableau VI.4: Hypothèses relatives aux conditions de réalisation

2c	Conditions de réalisation
2c1	Il existe un lien entre le délai de préparation imparti par l'établissement et la qualité du travail de maturité.
2c2	L'existence et la connaissance d'un guide se traduisent par une meilleure qualité des travaux de maturité.
2c3	La transparence perçue des modes opératoires a un effet positif sur la qualité des travaux de maturité.
2c4	La transparence perçue des critères d'évaluation a un effet positif sur la qualité des travaux de maturité.
2c5	Il n'existe aucune corrélation entre l'appréciation portée sur les critères d'évaluation et la qualité.

2.5.3 Niveau du suivi

Les élèves sont en général suivis par un ou plusieurs enseignants pendant la préparation de leur travail de maturité. Ces enseignants ont à cette occasion un impact direct ou indirect sur la qualité du travail produit. Les modalités du suivi ont des effets indirects sur elle, par exemple si l'enseignant éveille l'intérêt et la motivation des élèves. Mais on peut aussi penser à d'autres facteurs ayant des effets directs sur la qualité du travail au niveau de l'encadrement, par exemple si le suivi intervient directement dans le travail. Il conviendra donc notamment de vérifier si les travaux des élèves satisfaits de la façon dont ils ont été suivis présentent une meilleure qualité que ceux des élèves qui en sont mécontents (→ hypothèse 3a). Le présent projet examine en particulier dans quelle mesure la satisfaction éprouvée à l'égard du soutien reçu dans chaque domaine se reflète sur les divers aspects de la qualité (→ hypothèses 3b et 3c).

On admet aussi que les élèves qui choisissent eux-mêmes l'enseignant qui les suivra optent pour une personne qu'ils jugent compétente et utile, ce qui veut dire qu'ils seront satisfaits de la qualité de leur encadrement (→ hypothèse 3d).

Contrairement à l'idée selon laquelle la satisfaction éprouvée à l'écart de l'encadrement aurait un effet positif sur la qualité du travail de maturité, la présente étude fait l'hypothèse que la fréquence des contacts de suivi et la forme de l'encadrement (contacts personnels, par courriel, suivi individuel ou en groupe) n'a pas d'effets sur la qualité des travaux de maturité (→ hypothèse 3e).

Tableau VI.5: Hypothèses relatives au suivi des élèves

3	Suivi des élèves
3a	Mieux l'élève est suivi, meilleure est la qualité du travail de maturité.
3b	La satisfaction éprouvée à l'égard de l'assistance reçue dans la formulation de la problématique et sur le contenu du travail présente une corrélation positive avec la qualité du contenu du travail de maturité.
3c	La satisfaction éprouvée à l'égard de l'assistance reçue sur les bases formelles, la structuration et l'articulation du travail présente une corrélation positive avec la qualité formelle du travail de maturité.
3d	Plus l'élève est libre dans le choix de l'enseignant qui l'encadrera, mieux le suivi est ensuite apprécié.
3e	Il n'existe pas de lien entre la fréquence et la forme des contacts d'encadrement d'une part, et la qualité du travail de maturité de l'autre.

2.5.4 Niveau individuel

Les facteurs individuels ont toujours un impact direct sur la qualité du travail de maturité. Mais il ne faut pas oublier qu'ils sont souvent influencés par l'extérieur, et donc par le cadre institutionnel, et particulièrement aussi par le suivi.

Motivation et perception de soi

Il convient d'examiner si le temps investi dans le travail de maturité – qui peut dépendre du temps débloqué par l'école pour le préparer et du délai de remise qu'elle fixe – a des effets sur sa qualité (→ hypothèse 4a1).

Au niveau individuel, il faut aussi s'interroger sur l'appui que reçoit l'élève de son entourage personnel. Les études de Pagnossin et al. (2005) et de Binder et Feller-Länzlinger (2003, 2004) avaient montré que l'entourage ou des tiers peuvent avoir une influence sur le choix du sujet; on peut également penser que les élèves reçoivent une assistance de leur entourage à la préparation de leur travail de maturité. On est ainsi conduit à faire l'hypothèse que les travaux de maturité des élèves bénéficiant d'une aide non seulement de l'enseignant chargé de les suivre, mais aussi d'autres personnes, ont tendance à être meilleurs (→ hypothèse 4a2).

On peut enfin penser que les élèves qui possèdent une connaissance préalable du sujet de leur travail soumettent des travaux de maturité d'un contenu de meilleure qualité que ceux qui se lancent dans un sujet nouveau pour eux (→ hypothèse 4a3). Il faut également vérifier si les élèves qui se considèrent armés des compétences nécessaires ont aussi tendance à préparer de meilleurs travaux (→ hypothèse 4a4), et se demander si l'intérêt déclaré de l'élève se répercute sur la qualité de son travail de maturité (→ hypothèse 4a5).

Tableau VI.6: Hypothèses relatives à la motivation et à la perception de soi

4a	Motivation et perception de soi
4a1	Il existe une corrélation positive entre le temps que l'élève rapporte avoir consacré à son travail de maturité et la qualité de ce dernier.
4a2	Il existe une corrélation positive entre l'assistance reçue de l'entourage personnel (parents, frères et sœurs, camarades) et la qualité du travail de maturité.
4a3	La connaissance préalable du sujet a un impact positif sur la qualité de fond du travail de maturité.
4a4	Les compétences (connaissances préalables et autonomie) que pense posséder l'élève présentent une corrélation positive avec la qualité du travail.
4a5	L'intérêt déclaré <i>a posteriori</i> par l'élève présente une corrélation positive avec la qualité du travail.

Valeur intrinsèque et extrinsèque du travail

Il convient par ailleurs d'examiner s'il existe un lien entre la qualité du travail d'une part et l'appréciation de son utilité et des compétences acquises en vue de qualifications ultérieures de l'autre, l'hypothèse étant que l'élève qui produit un travail de bonne qualité portera aussi une appréciation favorable sur l'utilité de ce travail et les compétences qu'il a acquises dans sa préparation (→ hypothèses 4b1, 4b2).

Sachant que les élèves avaient déjà une idée de la qualité de leur travail de maturité par l'appréciation de l'enseignant au moment de l'enquête, on est conduit à penser que l'appréciation des compétences acquises est liée à l'appréciation émanant de l'enseignant (→ hypothèse 4b3).

Tableau VI.7: Hypothèses relatives à la valeur intrinsèque et extrinsèque du travail

4b	Valeur intrinsèque et extrinsèque du travail
4b1	Il existe une corrélation positive entre la qualité du travail et l'appréciation de son utilité dans l'obtention de futures qualifications.
4b2	Il existe une corrélation positive entre l'appréciation par l'élève de l'utilité du travail et celle de l'accroissement de ses propres compétences.
4b3	Il existe une corrélation positive entre l'appréciation du travail par l'enseignant et l'appréciation par l'élève des acquis ainsi obtenus.

2.6 Population étudiée et échantillonnage

2.6.1 Population étudiée

La population étudiée se composait de tous les élèves qui préparaient un travail de maturité au cours de l'année scolaire 2006-2007. La durée de la formation gymnasiale et les délais de préparation du travail de maturité variant (jusque parfois au sein d'un même établissement)⁴⁸, l'échantillonnage s'est fondé sur les données relatives à tous les élèves qui se trouvaient en première année de formation gymnasiale (10^e année scolaire) pendant l'année scolaire 2004-2005, cela dans l'idée que ces élèves terminaient leur formation gymnasiale au cours de l'année 2006-2007 (tableau VI.8). Les statistiques disponibles conduisent à estimer que la population

⁴⁸ Dans la plupart des écoles, le travail de maturité se prépare au cours de l'avant-dernière ou de la dernière année (Zillig, 2004, p. 18).

étudiée comportait ainsi quelque 21 980 élèves et 1 044 classes. En moyenne nationale, les classes avoisineraient donc les 21 élèves (Suisse alémanique: 20,6; Suisse romande: 22,1; Suisse italienne: 19,4).

Tableau VI.8: Population étudiée (élèves en première année de formation gymnasiale au cours de l'année scolaire 2004-2005)

	Suisse	CH alémanique	CH romande	CH italienne
Filles	12 357 (56,2%)	6 706 (56,5%)	4 834 (56,4%)	817 (53,4%)
Garçons	9 623 (43,8%)	5 166 (43,5%)	3 743 (43,6%)	714 (46,6%)
Total	21 980 (100%)	11 872 (54%)	8 577 (39%)	1 531 (7%)

(Source: OFS)

2.6.2 Echantillonnage

Le mode d'échantillonnage a repris la méthodologie adoptée dans EVAMAR I (cf. Ramseier, 2005). Les travaux de maturité et les élèves eux-mêmes n'ont pas fait l'objet d'un échantillonnage aléatoire, mais systématique. L'échantillon a été défini dans un premier temps à l'échelle de l'école. On s'est fondé pour cela sur la *Liste des écoles dont la Confédération et la Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP) reconnaissent les certificats de maturité* (état au 1^{er} mars 2005) que tient la Commission suisse de maturité. Cette liste est ventilée par cantons; et au sein de chaque canton, elle présente les écoles dans l'ordre alphabétique, après le chef-lieu. Elle comporte en tout 169 écoles. Après élimination de toutes les écoles de maturité pour adultes et des écoles suisses à l'étranger, il est resté 158 établissements à prendre en considération dans l'étude.

On a ensuite procédé à une stratification de l'échantillon en régions linguistiques, pour garantir la représentation convenable de ces dernières. Cette répartition de l'échantillon est d'abord nécessaire pour des raisons de logique de la recherche. Mais en outre, EVAMAR I avait montré que les conditions de réalisation du travail de maturité variaient d'une région linguistique à l'autre, particulièrement en ce qui concerne la liberté de choix du sujet du travail de maturité. Il n'a pas été jugé bon de pousser plus loin la stratification (en cantons, tailles d'établissements, etc.), ces paramètres n'étant pas censés avoir d'influence sur la qualité des travaux de maturité.

Au sein de la strate explicite, les écoles ont été triées de façon à rapprocher les établissements d'un même canton, ce qui s'est traduit par une stratification implicite. Les établissements ont ensuite été pris au hasard: on a retenu un établissement sur x , en choisissant x de façon à obtenir le nombre souhaité M d'écoles à partir du total N de la strate. Ce qui donne une meilleure idée des valeurs moyennes qu'un échantillonnage purement aléatoire (Lohr, 1999, cité dans Ramseier, 2005, p. 6). Dans cette méthode, la probabilité de sélection d'un établissement est de M/N .

Tableau VI.9: Population et échantillon de gymnases

Gymnases	Population	Echantillon
Suisse	158 (100%)	30 (100%)
Suisse alémanique	118 (74,4%)	16 (53,3%)
Suisse romande	35 (22,2%)	12 (40%)
Suisse italienne	5 (3,2%)	2 (6,7%)

Dans chaque école retenue a été sélectionné un échantillon de classes préparant le travail de maturité au cours de l'année 2006-2007. Dans les établissements ne comportant pas plus de deux classes, l'enquête a englobé tous les élèves; dans les autres, deux classes ont été tirées au sort⁴⁹. L'enquête auprès des élèves a englobé tous les élèves des classes retenues.

Il ressort clairement de ces explications que la probabilité que tel ou tel élève soit retenu dans l'échantillon dépend de la taille de l'établissement. Il ne s'agit donc pas d'un échantillon représentatif, les élèves des petits établissements étant mieux représentés que les autres. On a surtout opté pour cette méthode pour obtenir dans l'ensemble un nombre de cas permettant de tester la corrélation entre la perception des élèves et les cadres formels mis en place par l'établissement. Les données n'ont été corrigées à l'analyse que lorsque la comparaison explicite avec celles du projet EVAMAR I imposait de maintenir la comparabilité sociodémographique.

L'étude de Pagnossin et al. (2005) montre que trois quarts environ de l'ensemble des travaux de maturité ont un caractère théorique ou empirique; les autres types sont nettement plus rares. C'est pourquoi, sachant que l'examen d'activités extrascolaires, de projets socioculturels et de productions artistiques, audiovisuelles et matérielles susciterait des problèmes de comparaison entre les travaux, n'ont été retenus pour le tirage que les travaux de maturité à caractère théorique ou empirique.

Avant échantillonnage des travaux à évaluer, il a été demandé aux écoles de classer les travaux de maturité des classes retenues selon qu'ils étaient théoriques, empiriques ou d'un autre type. En pratique, il est vite apparu qu'il est impossible de procéder à une claire classification sur ce critère. Nos échantillons ont toujours comporté des travaux traitant de productions techniques ou créatives, ou de réunions ou spectacles, tout cela étant très souvent considéré comme des travaux empiriques par les écoles consultées.

Sur la base de cette liste et des retours de questionnaires d'élèves, nous avons enfin tiré au sort dans chaque classe de l'échantillon un certain nombre d'élèves pour lesquels nous disposions d'un questionnaire et d'un travail de maturité.

Cet échantillonnage systématique se traduit par une perte d'efficacité statistique. L'effet design donne la mesure de cette perte de précision et d'efficacité d'un échantillon complexe comme notre échantillon stratifié en grappes, par rapport à un échantillon aléatoire simple, et indique dans quelle mesure la taille effective de l'échantillon est réduite par son design. L'effet design dépend de l'homogénéité (corrélation intraclasse) et de la taille des grappes (établissements, classes): au sein d'une école ou d'une classe, les élèves s'influencent les uns les autres en ce qui concerne la qualité du travail de maturité, et s'expriment donc d'une façon plus homogène qu'au sein d'un groupe aléatoire. Pour tirer des conclusions fiables des données collectées, il faut agrandir en conséquence l'échantillon, pour arriver à une hétérogénéité représentative malgré l'effet design.

Pour cette étude, nous avons adopté une valeur de l'effet design $deff = 4,5$. A titre de comparaison, elle avait été fixée à 6,6 pour l'étude Civic-Education, et à 5,5 pour l'étude PISA 2000 mathématiques. Mais notre étude porte sur la qualité des travaux de maturité, où l'effet de classe sur la préparation du travail est bien moins important que dans ces deux enquêtes. L'organisation de la préparation du travail de maturité n'est en effet pas liée au contexte spécifique d'une classe constituant un groupe d'apprentissage en interaction avec certains enseignants; c'est pourquoi nous avons fait l'hypothèse que l'effet design est plus faible.

Pour déterminer la taille de l'échantillon, nous sommes partis du fait que la valeur moyenne de l'échantillon doit être de $\pm 0,2$ l'écart-type de la valeur moyenne de la population si l'on veut passer avec une certaine précision de l'échantillon à la population. Pour obtenir la précision suffisante (95%) dans la détermination des valeurs moyennes, il faut donc partir d'une taille d'échantillon de 100 environ⁵⁰. Compte tenu d'une valeur de l'effet design $deff = 4,5$, on obtient une taille d'échantillon approximative de $n=450$.

⁴⁹ Deux nombres aléatoires y_1 et y_2 ont été générés entre 1 et x (où x = nombre de classes dans l'école). On a ensuite pris les classes occupant les positions y_1 et y_2 dans la liste soumise par l'école.

⁵⁰ La taille nécessaire de l'échantillon dans les conditions décrites est donnée par la formule $n = (SD/SE)^2$, où SD est l'écart type de la population et SE l'erreur type de l'échantillon. Pour une précision de 95%, les conditions décrites conduisent à $1.96 SE = 0,2 SD$. D'où $n = 96$.

2.6.3 Collecte des données

Une fois les écoles échantillonnées, une lettre leur a été envoyée au début de l'année 2007 pour leur demander de prendre part à l'enquête. Sur les 30 établissements (16 en Suisse alémanique, 12 en Suisse romande, 2 en Suisse italienne), 13 ont dû être remplacés (7 en Suisse alémanique et 6 en Suisse romande) soit parce qu'ils n'offraient plus de formation gymnasiale (2), soit parce qu'ils ne souhaitaient pas s'associer au projet (11).

Les questionnaires destinés aux élèves ont été envoyés aux écoles entre les mois d'avril et de juin 2007. Sur le total de 1 073 élèves ainsi contactés, 920 ont répondu (542 garçons et 378 filles), soit un taux de retour de 85,74%. Les âges s'étaient entre 17 et 25 ans, la moyenne étant de 18,9 ans (SD=1,01) au moment de l'enquête.

Le taux de retour a varié entre les régions linguistiques. En Suisse romande, les questionnaires n'ont pu être remplis en classe que dans huit écoles, la plupart des cours normaux étant remplacés par des plages de préparation personnelle au dernier trimestre avant les épreuves de maturité. Il en est allé de même dans une école tessinoise, où nous avons directement écrit aux élèves quand c'était la seule solution. Il est rapidement apparu que le taux de retour était plus faible dans ces établissements. Mais dans ce contexte, il faut aussi tenir compte de la répartition des personnes interrogées entre les régions linguistiques: 59,3% d'entre elles pour les écoles de maturité alémanique, 34,7% pour les écoles romandes, et 6% pour les écoles tessinoises. Par rapport à la population totale, où les proportions correspondantes sont respectivement de 54%, 39% et 7%, l'échantillon est donc légèrement distordu, et il convient de relativiser en conséquence la portée des résultats régionaux.

Sur les 950 questionnaires d'élèves reçus, 437 ayant réalisé un travail de maturité théorique ou empirique ont été sélectionnés selon la méthode décrite ci-dessus⁵¹. Leurs travaux de maturité ont fait entre les mois de juin et de décembre 2007 l'objet d'une évaluation à laquelle ont procédé 42 évaluateurs au total en se fondant sur la grille de critères.

3 Cadres institutionnels

Cette section compare les cadres de préparation des travaux de maturité. Les objectifs et buts, les modalités d'organisation du travail, l'encadrement et l'évaluation ont lieu à divers niveaux selon le canton et l'école. Ces écarts s'expliquent par le système fédéraliste, qui veut que les cantons définissent eux-mêmes leur législation en la matière. Cette liberté a toutefois ses limites, elle peut être restreinte par la législation fédérale, ou céder devant des pressions à l'adaptation. L'analyse des conditions de réalisation des travaux commence donc par le plan d'ensemble national, pour descendre ensuite au plan moyen du canton, et envisager enfin le plan rapproché des établissements représentés dans l'échantillon⁵².

3.1 Directives générales

3.1.1 Directives nationales

A l'échelle nationale, c'est l'article 10 du RRM 95 qui fixe le cadre général de réalisation du travail de maturité:

Art. 10 Travail de maturité

Chaque élève doit effectuer, seul ou en équipe, un travail autonome d'une certaine importance. Ce travail fera l'objet d'un texte ou d'un commentaire rédigé et d'une présentation orale.

⁵¹ Le nombre effectif de travaux de maturité examinés est de $n = 406$. Il est un peu inférieur à celui des élèves consultés, ce qui s'explique par le fait que certains des élèves de notre échantillon avaient préparé leur travail de maturité ensemble.

⁵² Il n'a pas été tenu compte des directives et règlements entrés en vigueur après 2006, l'échantillon se référant aux directives de 2006.

Ces règles portent donc avant tout sur l'organisation du travail et ses instruments. Il ne mentionne pas explicitement les objectifs ni le statut du travail de maturité, ni ne fait référence à sa fonction de propédeutique scientifique ou à ses contenus possibles. En revanche, l'élément central est *l'autonomie*: cette notion renvoie à des formes de travail autres que les modèles pédagogiques usuels du gymnase, mais aussi à l'indépendance et à la créativité attendues des élèves dans leur travail de maturité.

Le travail de maturité est également mentionné dans les articles 15 et 20 du RRM 95. On y trouve d'une part que le travail de maturité est évalué sur la base des prestations écrites et orales, mais aussi que le certificat de maturité doit mentionner le titre et la note du travail de maturité. Cela dit, cette note n'a pas été prise en compte jusqu'à présent dans le certificat de fin d'études secondaires, et l'indication du sujet et de la note n'a donc que valeur symbolique. L'article 5 (objectif des études) du RRM 95 contient d'autres indications et exigences applicables à la préparation d'un travail de maturité, sans toutefois faire explicitement référence au travail de maturité lui-même; on trouve à l'al. 2 l'esquisse d'une fonction de propédeutique scientifique assignée à l'enseignement gymnasial «ils se familiariseront ainsi avec la méthodologie scientifique». Et en ce qui concerne la proportion de temps à lui consacrer, l'article 11 prévoit de 15% à 25% en tout pour l'option spécifique, l'option complémentaire et le travail de maturité.

A l'échelle nationale, il faut aussi évoquer le plan d'études cadre pour les écoles de maturité (CDIP, 1994), qui ne prescrit rien à proprement parler sur le contenu du travail, mais se contente de fixer des objectifs généraux valables pour tous les élèves. Il y est par exemple indiqué, à propos des orientations générales des études gymnasiales, que «au niveau instrumental, le gymnasien apprend à formuler des énoncés théoriques, à énoncer des hypothèses, à les vérifier ou les infirmer, à en tirer des conclusions» (CDIP, 1994, p. 11 ss.). L'élève doit en outre se familiariser avec les bases méthodologiques du travail scientifique. Mais il revient en général aux écoles de donner corps à ces contenus et de les concrétiser.

3.1.2 Objectifs cantonaux

Dans notre échantillon, les dispositions cantonales relatives au travail de maturité figurent à divers échelons de la loi⁵³. La plupart des cantons reprennent dans des lois ou ordonnances cantonales les prescriptions peu concrètes du RRM 95, qui peuvent à l'occasion être complétées par des dispositions plus détaillées. Les plans d'études cantonaux fournissent à ce niveau intermédiaire une autre source, qui précise parfois encore les objectifs.

Objectifs définis par les textes juridiques cantonaux

Parmi les cantons alémaniques représentés dans notre échantillon, une petite partie seulement ont défini pour toutes leurs écoles des objectifs du travail de maturité allant nettement au-delà des exigences formulées dans le RRM 95. Dans quatre d'entre eux (Bâle-Campagne, Bâle-Ville, Lucerne, Obwald), des textes définissent non seulement les objectifs, mais aussi de façon détaillée la forme, le déroulement et les résultats du travail de maturité. Ces règles détaillées présentent une homogénéité de base en ce qui concerne la propédeutique scientifique, c'est-à-dire le caractère méthodologique systématique de la préparation du travail de maturité, ainsi que l'autonomie dont doit faire preuve l'élève dans ce contexte; elles s'écartent ensuite nettement dans leur contenu comme dans le degré de détail de leur formulation.

⁵³ 18 cantons, c'est-à-dire: Argovie, Bâle-Campagne, Bâle-Ville, Berne, Fribourg, Genève, Grisons, Jura, Lucerne, Neuchâtel, Obwald, Saint-Gall, Tessin, Thurgovie, Vaud, Valais, Zoug, Zurich.

Exemple 1: Règlement du 2 février 2001 relatif à la préparation des travaux de maturité dans les gymnases du canton de Bâle-Ville

§ 1. Buts

(1) Les élèves doivent, pour leur travail de maturité, formuler sur un sujet une tâche complexe, l'analyser et lui trouver des solutions. Ils doivent recourir pour cela à des méthodes et à des outils appropriés.

(2) Les élèves peuvent en particulier:

- chercher par eux-mêmes un sujet, le délimiter, définir des questions intéressantes qu'il soulève, puis le fixer conformément aux dispositions du § 10, en accord avec un enseignant de leur établissement;
- se procurer par eux-mêmes l'information nécessaire;
- aborder cette matière de façon différenciée;
- présenter le sujet dans le respect des impératifs de forme, de construction logique et de langue;
- concevoir un objet et son devenir sur le mode cognitif;
- présenter et expliquer leur propre travail à l'oral;
- se montrer capables de travailler avec des spécialistes dans un but précis;
- examiner des questions pluridisciplinaires si le sujet s'y prête;
- traiter un sujet pendant une période prolongée en faisant preuve de curiosité et d'intérêt.

Dans la plupart des cantons, le législateur s'est borné à répéter les dispositions de l'article 10 du RRM 95 et à mentionner les conditions à remplir pour passer avec succès les épreuves.

Exemple 2: Ordonnance du 6 juillet 1999 relative aux études gymnasiales du canton des Grisons.

Art.16 Travail de maturité

(1) Chaque élève doit effectuer, seul ou en équipe, un travail autonome, d'une certaine importance et conforme au Règlement sur la reconnaissance des certificats de maturité gymnasiale. Ce travail fera l'objet d'un texte ou d'un commentaire rédigé et d'une présentation orale.

(2) Les travaux écrits et oraux effectués à l'occasion du travail de maturité seront évalués.

(3) Dans le certificat de maturité, la note obtenue pour le travail de maturité sera exprimée par les mentions excellent, très bon, bon, satisfaisant, suffisant ou insuffisant.

Il en va de même en Suisse romande et au Tessin. Dans les cantons de Fribourg, de Genève, du Jura, de Neuchâtel, du Tessin, de Vaud et du Valais, il existe des textes juridiques applicables à toutes les écoles, auxquels les établissements se réfèrent plus ou moins directement dans leurs directives, mais qui ne vont pas sensiblement au-delà des directives du RRM.

En résumé, le travail de maturité figure dans la législation de tous les cantons, mais à divers niveaux. Cela pourrait montrer que les législations cantonales accordent une priorité variable à la définition des objectifs de formation. Les textes adoptés par le Parlement jouissent d'une légitimité démocratique plus grande, d'une assise plus large, et d'une prééminence formelle sur les ordonnances et décrets émanant de l'exécutif. Mais il ne faut pas oublier non plus que les ordonnances se révèlent plus souples et plus ajustables, s'agissant d'un dispositif décisionnel plus léger. Et l'on ne peut s'empêcher de remarquer que cet ancrage légal est lui-même aussi soumis au processus législatif, et donc dans une mesure notable aux contingences de la vie politique⁵⁴. Il revient la plupart du temps à l'école elle-même de définir plus précisément les objectifs du travail de maturité et de concrétiser leur réalisation dans des règlements et des directives.

Dispositions figurant dans les plans d'études cadres cantonaux

Les plans d'études cadrent cantonaux ne contiennent que rarement de consignes applicables à toutes les écoles sur le travail de maturité. Celui du canton d'Argovie, par exemple, mentionne l'organisation et les buts du travail de maturité; celui du canton de Berne, le but général du travail de maturité, le détail de l'exécution étant confié aux gymnases; ceux des cantons de Bâle-Ville et de Thurgovie contiennent aussi des observations introduisant le

⁵⁴ L'interprétation exacte de ce phénomène imposerait d'examiner plus précisément la documentation du processus législatif, ce qui dépasserait le cadre de la présente analyse.

travail de maturité. Ceux de la plupart des cantons ne mentionnent même pas le travail de maturité, ou ne le font que pour indiquer que l'interprétation en revient à chaque établissement.

3.2 Directives émanant des établissements

L'analyse repose sur des informations à caractère impératif formulées et documentées par les écoles⁵⁵. L'établissement a surtout été considéré ici dans sa fonction d'organisation, sachant bien que d'autres explications et conseils sont fournis par les professeurs de diverses matières ou dans des cours spécifiques, mais qu'il n'est pas possible d'en tenir compte dans une étude nationale. Les documents dont on dispose ici représentent la base la plus utile d'analyse qualitative, car ce sont eux qui concrétisent les exigences que s'impose à elle-même l'école dans sa fonction gymnasiale, et lui servent de carte de visite à l'intérieur comme à l'extérieur.

3.2.1 Attributions et directives

Il convient d'abord d'indiquer qu'une petite moitié seulement (43%) des écoles étudiées se sont dotées de règles spécifiques au travail de maturité. Il s'agit le plus souvent de règlements généraux, de règlements de passage ou de directives adoptées officiellement – en général par la commission scolaire du gymnase. Les textes spécifiques émanant de l'école sont rares en Suisse latine, mais pas en Suisse alémanique.

Il existe dans 30 des gymnases étudiés une commission ou un groupe de travail spécifiquement chargé de la coordination et du déroulement des travaux de maturité pour l'ensemble de l'école; ces organes sont formés le plus souvent d'enseignants et de membres de la direction de l'école. Les autres établissements ne mentionnent pas d'organes de ce type. Dans tous les gymnases de Suisse latine (à l'exception des écoles cantonales du canton de Fribourg), c'est la direction de l'école qui assume la responsabilité de l'exécution.

Tous les gymnases sans exception ont préparé des guides ou des directives censés donner des renseignements sur l'organisation, le travail et les résultats attendus. Le contenu de ces documents varie considérablement, allant d'une information rudimentaire sur l'idée générale, l'encadrement et la notation, jusqu'à des recueils d'informations très détaillées couvrant des sujets généraux, mais aussi des consignes spécifiques à chaque matière. La plupart des établissements préparent des documents plus ou moins détaillés abordant *l'organisation du travail, le choix du sujet, les exigences formelles, la présentation et l'évaluation*, ainsi que le calendrier.

Les écoles cantonales de Suisse romande et du Tessin fournissent aussi de l'information liée à chaque matière. Cela découle de l'organisation en séminaires où des enseignants approfondissent un sujet général dans des séances spéciales avec un petit cercle d'élèves. Compte tenu du rôle primordial que joue l'enseignant assurant cet encadrement dans cette forme d'organisation, le degré de concrétisation de l'information fournie par écrit n'est pas très élevé dans ces gymnases.

L'article 10 du RRM 95 prévoit que le travail de maturité peut être préparé seul ou en équipe. Sept écoles limitent la taille des équipes à deux ou trois personnes; les autres ne mentionnent pas de chiffre précis.

3.2.2 Objectifs définis par les établissements

Presque tous les établissements formulent les objectifs que doivent atteindre les élèves dans la préparation de leur travail de maturité. En général, il s'agit de se familiariser avec la réflexion scientifique et la méthodologie en se concentrant sur le travail de maturité gymnasiale. On retrouve presque partout aussi le travail sur un projet, faisant une large place à la capacité de s'organiser. Dans la liste ci-dessous, on remarque que la majorité des écoles alémaniques étudiées n'accordent aucune priorité au travail interdisciplinaire, qui figure pourtant

⁵⁵ L'anonymat ayant été promis aux écoles, elles ne seront pas nommées dans la suite, mais simplement classées selon leur région linguistique.

explicitement dans les buts formulés par presque toutes les écoles de Suisse romande. De même, l'accumulation de savoirs figure uniquement dans les objectifs définis par les écoles romandes (tableau VI.10).

Tableau VI.10: Objectifs, 2006 (N=30)

	Autonomie (indépendance)	Communication (présentation écrite et orale)	Méthodes & réflexion scienti- fique	Techniques d'organisation et gestion des délais	Créativité, origina- lité	Interdisciplinarité	Accumulation de savoirs
D CH 1	■	■	■				
D CH 2	■		■		■		
D CH 3	■	■	■	■	■		
D CH 4							
D CH 5	■	■	■	■			
D CH 6							
D CH 7	■		■	■			
D CH 8			■		■		
D CH 9	■			■			
D CH 10	■	■	■	■			
D CH 11							
D CH 12	■	■	■	■			
D CH 13	■	■	■	■			
D CH 14	■	■	■	■			
D CH 15	■	■	■	■	■		
D CH 16			■	■	■		
F CH 1		■	■	■	■		
F CH 2	■	■	■	■	■	■	
F CH 3		■	■				
F CH 4		■	■	■	■	■	■
F CH 5							
F CH 6	■	■	■	■	■		
F CH 7	■	■	■			■	■
F CH 8	■	■	■	■	■	■	■
F CH 9	■	■	■			■	■
F CH 10	■	■	■			■	■
F CH 11			■	■	■	■	
F CH 12	■	■	■	■	■	■	■
I CH 1		■	■			■	
I CH 2			■	■			
Total mentions	21	19	25	18	12	9	6

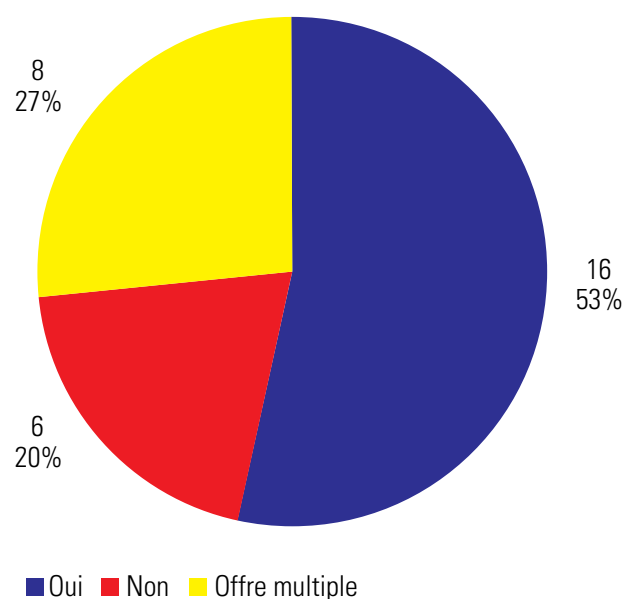
Il faut souligner dans ce contexte que l'autonomie figurant dans le RRM 95 reparait dans certains objectifs sous forme d'*indépendance*, avec un certain décalage sémantique. La distinction à première vue minime (les deux

notions se référant en principe à la capacité de réaliser le travail par soi-même) peut toutefois se traduire par des différences sensibles dans la part de l'apport créatif de l'élève. La réflexion autonome s'appuie sur la recherche et le traitement indépendants de l'information, sans se ramener à eux⁵⁶. On notera aussi que diverses écoles alémaniques (5 sur 16) mentionnent explicitement la créativité et l'originalité à propos du produit final. On trouve par exemple que le travail de maturité doit être «significatif», ou qu'il doit être «original, voire inédit par certains aspects notables». On remarque aussi une autre particularité, qui met l'accent sur le produit du travail de maturité; une école cantonale veut par exemple que ce soit «une représentation possédant une visibilité extérieure, à valeur de communication et revêtant une forme ambitieuse».

3.2.3 Préparation et choix du sujet

Sur les 30 gymnases étudiés, 24 prévoient des préparations spécifiques au travail de maturité (figure VI.4). Il s'agit souvent de séances d'information réunissant plusieurs classes, où est diffusée une information sur les buts, les modalités du travail et le calendrier. On observe de très grosses différences d'une école à l'autre en ce qui concerne les préparations allant au-delà, et sur le choix du sujet. Huit des trente écoles de notre échantillon organisent d'autres cours: les élèves peuvent par exemple suivre un cours de rédaction, en plus de la séance d'information générale; dans un cas, il est aussi possible de s'exercer par des travaux préalables et dans des unités d'enseignement aux méthodes applicables à un projet, avant d'enrichir cette première pratique au cours du travail de maturité.

Figure VI.4: Séances spécifiques de préparation au travail de maturité dans les gymnases

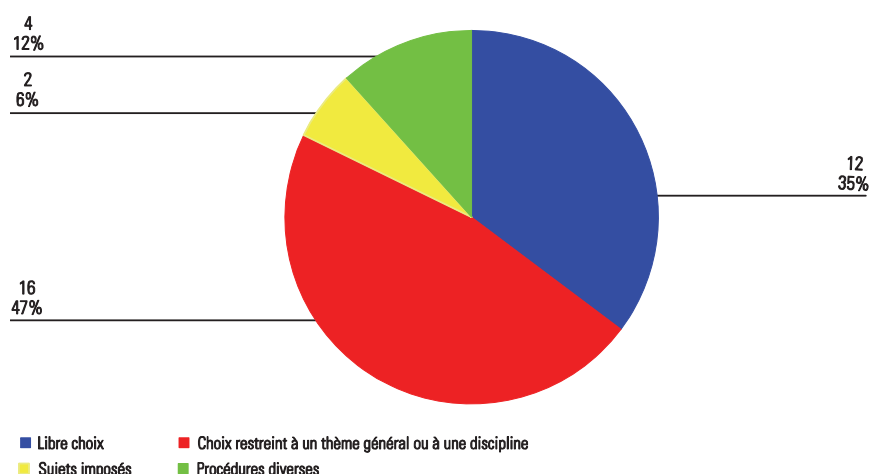


Les écoles cantonales de Suisse latine organisent surtout des séminaires de préparation, traitant d'un sujet général fixé d'avance – le plus souvent par l'enseignant. Une quinzaine d'élèves peuvent s'y inscrire autour de chaque enseignant, avant de choisir leur sujet. Cette organisation thématique du travail limite les sujets possibles: une palette de contenus est ainsi définie, les élèves y choisissant l'objet qu'ils veulent traiter.

En ce qui concerne le choix du sujet, on note aussi que les gymnases laissent une grande liberté à leurs élèves. Un tiers des écoles étudiées environ ne demande même pas que le travail soit lié à une matière, ce qui veut dire qu'il ne pourra peut-être pas être suivi par un enseignant spécialiste du domaine correspondant. Une petite moitié des gymnases insiste pour que le travail de maturité soit en rapport avec une ou plusieurs matières.

⁵⁶ Bonati & Hadorn (2007, p. 17) ont fait la même constatation.

Figure VI.5: Modalités de choix du sujet (N=30)



La possibilité d'opter pour un sujet interdisciplinaire n'est exclue par aucune des écoles de notre échantillon; elle ne figure toutefois explicitement que dans les directives des écoles romandes (cf. 3.2.2). Une minorité de gymnases admettent plusieurs procédures: libre choix, libre choix dans une matière, choix dirigé par l'enseignant. Dans toutes les écoles examinées, l'élève soumet son sujet à l'approbation de l'enseignant assurant son accompagnement avant de commencer à y travailler.

3.2.4 Consignes formelles

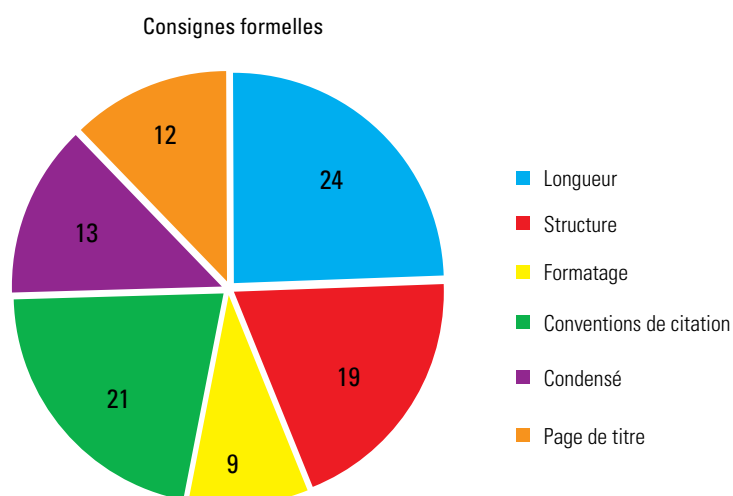
La plupart des écoles ont des directives formelles sur le travail de maturité, avec diverses règles sur la longueur du texte, sa structure et sa présentation (figure VI.6). Les consignes formelles concernent le plus fréquemment la longueur du texte; c'est le cas dans 24 de nos 30 établissements, avec de fortes variations (de 10 à 30 pages)⁵⁷. Trois établissements (tous alémaniques) indiquent en outre que la longueur minimum du texte dépend des consignes en vigueur dans la discipline concernée; six ne donnent aucune indication de longueur.

Les règles de présentation des citations et de la bibliographie sont également expliquées dans le détail (21 mentions). Les écoles donnent par ailleurs, le plus souvent sous forme d'exemples, des instructions sur les façons et les raisons de citer et de donner les références bibliographiques (lutte contre le plagiat). Plus de la moitié d'entre elles abordent aussi la structure générale du rapport (19 mentions). On constate toutefois des écarts considérables à ce niveau, depuis des indications très générales (page de titre, introduction, corps, conclusion) jusqu'à des plans pratiquement complets.

On observe dans ce contexte des différences entre régions linguistiques: les consignes de formatage n'apparaissent que dans les directives des gymnases alémaniques, le spectre allant de la force des caractères et de la présentation du titre jusqu'à des exemples détaillés de présentation (alignement du texte, procédés de mise en relief, formatage des notes de bas de page, etc.). Le condensé ou la synthèse précédant le travail à proprement parler sont surtout obligatoires en Suisse romande et au Tessin.

⁵⁷ Le chiffre indicatif est en général plus élevé pour les travaux de groupe.

Figure VI.6: Consignes de forme (nombre de mentions)

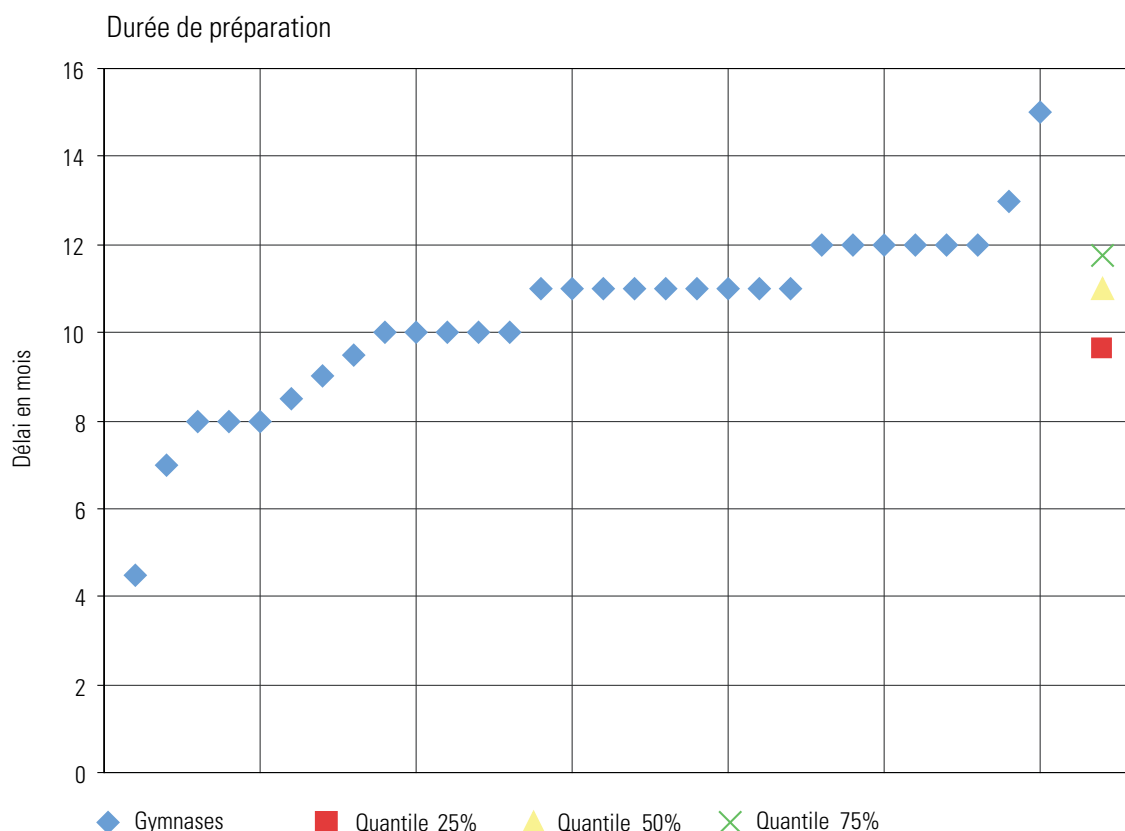


3.2.5 Mode opératoire et calendrier

Dans toutes les écoles sauf une, le travail de maturité se prépare au cours de l'avant-dernière ou de la dernière année avant la maturité. Le travail s'organise en général de la façon suivante: après une introduction (dans un cours normal ou en séance spéciale, cf. 3.2.3 ci-dessus), commence la recherche d'un sujet et de l'enseignant qui suivra le travail. Cette première phase prend le plus souvent plusieurs mois, mais peut aussi ne durer que quelques semaines. La deuxième phase est occupée par la collecte et le dépouillement de l'information, la préparation d'un plan détaillé et, en règle générale, la rédaction d'un premier jet du rapport; les écoles prévoient de quatre à huit mois pour elle. Le travail de maturité est le plus souvent rédigé dans les deux mois précédant la remise du document. Partout sauf dans deux établissements, la présentation orale a lieu après la remise du travail (parfois tout de suite après, mais jusqu'à deux mois plus tard).

Le délai imparti aux élèves entre le choix définitif de leur sujet et la remise du travail varie dans les extrêmes de 4,5 à 14,5 mois; mais mis à part ces cas exceptionnels, la plupart des écoles accordent entre neuf (env. 25%) et onze mois de préparation (env. 75% des gymnases étudiés).

Figure VI.7: Variance du délai de préparation (mois)



Dans la moitié environ des établissements alémaniques, les élèves disposent d'un appoint de temps d'enseignement consacré à la préparation du travail (au lieu des cours normaux, en plus de la réduction du nombre d'heures dans l'horaire général des classes de maturité). Les choses sont organisées différemment en Suisse latine: la formule du séminaire se traduit automatiquement par des plages d'enseignement consacrées à la préparation du travail.

Une grande partie du travail se fait toutefois en dehors des cours. La plupart des écoles ne spécifient pas le temps à lui consacrer: quatre seulement des 30 établissements étudiés donnent des indications approximatives pour la rédaction et la présentation (entre 80 et 100 heures de travail). Mais on constate que dans l'ensemble, le temps de travail des élèves se répartit sur une longue période. On ne constate que rarement de plages prolongées de travail intensif.

3.2.6 Suivi du travail de maturité

Un grand nombre de gymnases alémaniques indiquent qu'un contrat de travail est conclu entre l'élève et l'enseignant qui l'accompagne. Ce document mentionne surtout le sujet et le plan général, ainsi que parfois la motivation et la forme de la coopération entre l'élève et l'enseignant. Il sert de base à la suite du travail, et dans certains cas de référence à l'évaluation ultérieure du travail de l'élève. Il ne s'agit pas du journal de travail qui sert dans toutes les écoles d'auxiliaire d'accompagnement et de progression, et qui témoigne et rend compte de toutes les étapes du travail. Dans quelques établissements, ce compte rendu documenté constitue une partie obligatoire du travail de maturité. La plupart des écoles tiennent compte de la réflexion sur le déroulement du travail dans l'appréciation finale du travail de maturité.

Le suivi est normalement assuré par un ou plusieurs enseignants, qu'il incombe le plus souvent à l'élève de trouver. Dans un nombre restreint d'établissements, il est possible de se faire encadrer ou co-encadrer par un expert extérieur. Plus rarement, un membre de la direction de l'école joue le rôle d'expert dans l'évaluation

finale. Les demandes d'encadrement, le plus souvent présentées sous forme de description écrite détaillée du champ d'investigation, sont examinées par les enseignants, et soumises à leur approbation explicite. On note à ce sujet que la compétence technique ne figure que dans de rares documents comme condition à remplir impérativement par la personne susceptible d'assurer le suivi. On peut penser qu'au-delà de la discipline et de la disponibilité, les affinités personnelles entre élèves et enseignants jouent aussi un rôle dans la sélection. Le rythme du suivi et la fréquence minimum des échanges ne sont explicitement mentionnés que dans un petit nombre d'écoles (9).

Le cahier des charges de l'enseignant assurant le suivi est très similaire entre les écoles étudiées. Il doit surtout assister l'élève sur le contenu du travail et au niveau méthodologique. Ce rôle est toutefois interprété de façon très variable, et va du conseil et de l'appréciation jusqu'à la surveillance et au contrôle disciplinaire, en passant par l'aide à «l'accouchement intellectuel». La plupart des écoles adoptent une position médiane, à mi-chemin entre l'accompagnement formateur et la simple homologation du travail.

Une petite partie seulement des gymnases alémaniques prévoient une initiation au suivi des travaux de maturité (réunion d'information ou documentation); aucun document ne mentionne de préparation spécifique dans les gymnases romands et tessinois. L'enseignant dispose comme l'élève du guide sur le travail de maturité, qui décrit plus ou moins concrètement les attentes auxquelles doit répondre le travail de maturité et les objectifs qu'il doit remplir.

On peut penser que les écoles estiment d'habitude que la formation générale des enseignants leur permet d'encadrer convenablement un travail de maturité. Quelques écoles seulement leur demandent explicitement de posséder des compétences professionnelles dans la matière dont relève le travail de maturité à encadrer; dans les cas limites, cela risque toutefois de se refléter sur la qualité de l'assistance fournie sur le fond.

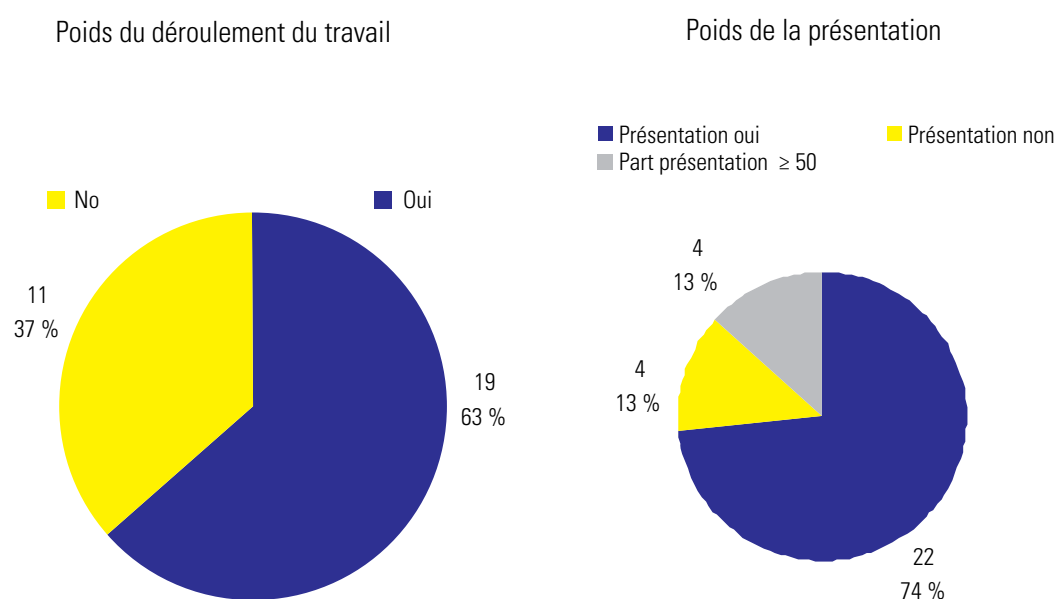
Le suivi des travaux de maturité est le plus souvent considéré comme faisant partie du service des enseignants, avec un maximum de trois à huit travaux à suivre. Seules huit écoles prévoient explicitement une décharge pour le suivi dans leur règlement; elle est en général donnée sous forme de diminution du nombre d'heures d'enseignement, ou par inscription dans la feuille de service.

3.2.7 Evaluation et notation

En ce qui concerne les critères d'évaluation, on peut distinguer trois variantes. Une partie des établissements (dont la moitié des gymnases alémaniques) prévoit des critères généraux applicables aux travaux de maturité de toutes les disciplines: les consignes contiennent une brève description des critères généraux d'évaluation de la qualité de l'idée, du projet, de sa mise en œuvre et de la présentation. Dans un deuxième groupe, les critères sont concrétisés par discipline; cela donne des schémas d'évaluation détaillée, débouchant sur des points concrets à apprécier. Dans ces deux groupes, les critères d'évaluation et de notation sont très similaires d'une école à l'autre en ce qui concerne les catégories couvertes: forme et expression, contenu et argumentation, approche méthodologique. Dans un troisième groupe d'établissements (dont la plupart en Suisse romande et au Tessin), il revient à l'enseignant assurant le suivi de définir les critères d'évaluation; il n'est pas spécifié si et dans quelle mesure les élèves ont connaissance de ces critères fixés par des spécialistes ou des enseignants.

On observe d'autres écarts de notation entre les écoles au niveau de la pondération des étapes et des produits du travail (figure VI.8). Ces différences de coefficients se rencontrent entre le travail écrit ou la documentation et la présentation orale. Cette dernière peut former jusqu'aux deux tiers de la note générale.

Figure VI.8: Pondération de la notation finale



La plupart des écoles indiquent que la notation englobe aussi le déroulement du travail. La majorité d'entre elles (63%) s'appuient par exemple sur la motivation et l'attitude devant le travail, le respect des délais, l'autonomie d'apprentissage. Il est parfois tenu compte également de l'auto-évaluation de l'élève. Et quatre établissements mentionnent explicitement la conception générale du projet.

Pour le travail d'équipe, la plupart des écoles exigent que l'on puisse discerner dans le produit final la part de chaque membre, et l'attribuer à son auteur. Il peut aussi arriver en revanche que le produit écrit soit évalué comme travail d'équipe, après accord, tandis que les présentations sont en général évaluées individuellement.

Une note insuffisante à des conséquences variables: dans la grande majorité des établissements, l'élève doit obtenir une note suffisante à son travail de maturité et à sa présentation pour être admis à se présenter aux épreuves de maturité. Dans une part relativement restreinte des gymnases, la présentation orale n'est possible qu'après obtention d'une note suffisante au travail écrit; dans d'autres établissements, la présentation orale est automatique.

Le RRM 95 stipule que le titre et la note doivent figurer dans le certificat, mais sans impact de la note sur l'obtention du diplôme. Dans la pratique, on trouve plusieurs formules: en général, le certificat mentionne le sujet et l'appréciation. Dans quelques écoles cantonales, la note du travail faisant l'objet du document est intégrée à la note obtenue dans une discipline (suivant celle dont relève le travail). Les gymnases surtout romands n'utilisent pas de système de notation à proprement parler, mais des appréciations.

3.3 Bilan partiel

Il existe de nombreux documents, certains très complets, permettant d'apprécier les conditions dans lesquelles sont préparés les travaux de maturité. De nombreuses écoles sont par ailleurs en train de préparer, d'affiner ou de tester leurs directives. Ces changements rendent un état des lieux difficile, mais pas impossible. La présente analyse des cadres documentés⁵⁸ s'appuie, comme pour l'évaluation des travaux de maturité, sur les consignes en vigueur dans les écoles pour l'année 2006-2007.

⁵⁸ Rappelons une fois encore que les écoles ne font pas figurer tous leurs accords de coopération dans des documents. Les ententes générales informelles de ce type échappent à notre analyse.

En résumé, on peut dire que la documentation fixant les attributions, les cadres généraux et les modes opératoires varie beaucoup d'un établissement à l'autre. La diversité des objectifs et des modalités de réalisation d'une tâche apparemment simple (voir 3.1.1 directives nationales) porte sur la forme (les éléments réglementés) et le contenu (formulation des objectifs et des consignes). Un certain nombre d'observations paraissent primordiales dans ce contexte.

- Il existe des **directives** à l'intention des élèves dans toutes les écoles étudiées. Mais la variabilité de leur contenu conduit à penser qu'elles n'ont pas partout le même statut: certaines, relativement allégées, laissent une marge de manœuvre importante à l'enseignant suivant le travail comme à l'élève dans le paramétrage spécifique du projet. Les descriptifs et exemples détaillés fournissent en revanche une aide concrète aux élèves et déchargent l'enseignant d'une partie du contrôle formel. Des exigences objectives favorisent la transparence des procédures et fixent les rôles de l'élève et de l'enseignant.
- On constate des écarts entre régions linguistiques au niveau de **l'organisation**: les gymnases des cantons de Fribourg, du Tessin, de Vaud et du Valais donnent à l'enseignement lié au travail de maturité la forme de séminaires, qui permet d'aborder des sujets sortant normalement des matières gymnasiales, et ouvre donc sur le cadre général de l'interdisciplinarité. Les travaux de maturité ainsi préparés traitent d'un sujet lié à une matière, mais sans être vraiment limités par elle. En Suisse alémanique, l'organisation mise d'habitude sur le principe de la motivation individuelle maximale des élèves, qui choisissent en général librement leur sujet et l'enseignant qui les suivra. Aucune de ces formes d'organisation n'implique que l'enseignant suivant le travail doive posséder de compétences spécifiques ou ne restreint le choix du sujet. Le choix de telle ou telle formule reflète donc moins une attitude pédagogique que la façon dont l'école affecte ses ressources.
- Presque toutes les documentations mentionnent l'initiation au **travail scientifique**. Il ne fait donc aucun doute que le travail de maturité a une fonction de propédeutique scientifique et simule le mode de travail des établissements d'enseignement supérieur. Il est alors d'autant plus étonnant que très peu d'écoles prévoient que la personne suivant le travail doive posséder une formation spécifique dans le domaine dont relève le travail: à s'en tenir aux prescriptions des écoles, il n'est en tout cas pas exclu que n'importe quel enseignant puisse s'occuper de n'importe quel travail. Les documents dont nous disposons ne nous ont pas permis de déterminer si le professionnalisme de l'encadrement figure dans des conventions non documentées.
- En ce qui concerne la **mesure des résultats**, on constate que contrairement à ce qui se fait à l'école primaire, les pratiques d'évaluation des travaux de maturité n'attirent guère l'attention. Le mode de notation varie donc énormément. L'évaluation synthétique peut englober des éléments formateurs dans certains établissements, et la présentation représente parfois plus de 50% de la note finale. On en vient à se demander comment est exactement notée la fonction de propédeutique scientifique: sur la base du zèle et de la motivation, des progrès atteints au cours du travail, du document écrit ou de la présentation orale? Et l'on ne sait encore pas si les changements décidés par la Confédération et les cantons en ce qui concerne le statut à accorder au travail de maturité clarifieront suffisamment les choses.
- En ce qui concerne les écarts et les similitudes notées au **niveau formel**, on remarque que les cantons ne peuvent pas s'appuyer sur une norme nationale uniforme, et qu'ils ne font rien eux-mêmes pour rattraper ou compenser cette lacune. Presque aucun canton de notre échantillon ne possède de directives uniformes, valables pour toutes les écoles, définissant les buts, les contenus et le déroulement du travail de maturité. Ce qui laisse aux écoles une grande latitude pour organiser elles-mêmes leurs modes opératoires, voire ne pas fixer de règles écrites. Cette absence de documents de référence dans certaines écoles ne doit pas amener à conclure à l'absence de règles: on peut faire l'hypothèse que bien des consignes sont données «sur le tas», c'est-à-dire en cours. Mais ces écarts formels considérables montrent clairement que le besoin de réglementer la préparation du travail de maturité ne fait pas l'unanimité, pas plus d'ailleurs à l'échelon national que cantonal. Cela laisse donc une liberté très ample aux établissements et aux enseignants pour

définir les conditions générales de réalisation des travaux; l'usage qu'ils font de cette liberté peut se traduire par de grandes différences dans les travaux de maturité, s'agissant d'un apprentissage dirigé. Il faudrait donc conduire à ce sujet une analyse longitudinale des travaux de maturité pour déterminer si la densité de réglementation de tel ou tel établissement lisse les variations de valeur des travaux au sein d'une classe terminale ou d'une discipline, voire d'une année à l'autre ou d'un canton à l'autre. Bien sûr, les directives écrites n'ont de sens que si elles ne restent pas lettre morte: pour qu'elles uniformisent les conditions de réalisation des travaux de maturité au sein d'une école, encore faut-il qu'elles soient appliquées.

4 Le travail de maturité vu par les élèves

L'enquête auprès des élèves est un volet important de l'analyse des travaux de maturité. Les réponses nous donnent des informations sociodémographiques de base sur notre échantillon, et nous renseignent par ailleurs sur le regard subjectif que portent les élèves sur divers aspects de leurs projets de fin d'études gymnasiales.

EVAMAR I avait déjà interrogé des élèves. Bien que l'étude ait eu un centrage plutôt systémique, le sous-projet 2 sur les objectifs pédagogiques transversaux avait posé des questions sur le travail de maturité, notamment en ce qui concerne *le choix du sujet, le suivi et les conditions de réalisation du travail*. C'est pourquoi il a été fait référence à EVAMAR I quand cela était possible dans le choix des questions, de façon à donner une dimension longitudinale à l'étude. Des items ont été repris sans changement ou moyennant de légers ajustements seulement en ce qui concerne le lien avec une discipline, le choix du sujet, le suivi et les conditions de réalisation.

Conformément au schéma de fonctionnement présenté en 2.1, les élèves ont été interrogés sur des facteurs individuels ayant un impact sur la qualité de leur travail de maturité. Il pouvait s'agir d'un retour sur la genèse, le déroulement et l'évaluation scolaire du travail, mais aussi de questions prospectives sur l'utilité des acquis dans d'autres processus qualifiants et dans la perspective du développement personnel. Les rubriques étaient les suivantes:

- renseignements sociodémographiques;
- matière(s) dont relevait le travail de maturité;
- choix du sujet;
- évaluation du travail de maturité;
- encadrement;
- assistance et ressources;
- appréciation du volume de travail consacré au travail de maturité et de son utilité.

Les sections suivantes reviennent sur ces aspects et présentent les résultats de l'enquête sous forme descriptive et avec une première analyse, qui sera approfondie dans les sections suivantes, à la lumière de l'évaluation des travaux.

4.1 Choix du sujet

Les options spécifiques des élèves relevaient à 31% de la langue première ou seconde, à 21,9% des sciences de la nature, à 31,7% des sciences humaines et sociales, et à 15,4% des arts, de la musique ou du sport. On observe de légères différences dans le choix des options par rapport aux résultats de l'étude EVAMAR I: parmi les élèves interrogés alors, 36% avaient choisi leur langue première ou seconde en option spécifique, 27% les sciences humaines et sociales, 25,8% les sciences de la nature et 11,1% un sujet relevant des arts, de la musique ou du

sport (Ramseier et al. 2005, p. 67)⁵⁹. Après correction, cette répartition devient la suivante: 31,6% pour la langue première ou seconde, 23,3% pour les sciences de la nature, 30,8% pour les sciences humaines et sociales, 14,3% pour les arts, la musique et le sport. On constate donc que les choix d'options spécifiques ont légèrement privilégié dans notre échantillon les arts appliqués et l'éducation physique ainsi que les sciences humaines et sociales, et que la langue première et seconde et les sciences de la nature ont été moins souvent retenues. La répartition est quelque peu différente pour l'option complémentaire: les sciences humaines et sociales accusent une nette avance, avec 54,7%, tandis que les sciences de la nature sont choisies dans 16,9% des cas, et l'art, la musique et le sport dans 28,4%.

4.2 Rattachement du travail de maturité à une matière et choix du sujet

À propos du choix du sujet, il est intéressant de s'interroger sur le rattachement du travail de maturité à une ou plusieurs matières. La question a donc porté sur le lien avec une matière principale et une matière secondaire, les réponses multiples étant explicitement admises dans les deux cas.

La plupart des élèves (66,4%) lient leur travail à une seule matière principale. Mais certains en mentionnent jusqu'à quatre. Au total, 20,8% des élèves nomment au moins deux matières principales, et 12,8% n'en mentionnent aucune; 7,9% ne rattachent leur travail de maturité ni à une matière principale, ni à une matière secondaire, et 49,9% ne situent leur travail que dans une matière principale ou une matière secondaire. Cela inclut les cas où la même matière est donnée comme principale et secondaire. On peut en conclure que dans l'optique des élèves, un peu moins de 50% des travaux sont monodisciplinaires; 28% d'entre eux trouvent que leur travail de maturité se rattache à deux matières, qu'elles soient qualifiées de principales ou secondaires; et 14,1% estiment que leur travail relève de plus de deux matières principales ou secondaires.

Ces résultats diffèrent sensiblement de ceux de la première étude EVAMAR, dans laquelle il avait été demandé aux élèves si leur travail de maturité concernait *une seule discipline scolaire, plusieurs, ou s'il s'agissait d'un sujet situé en dehors des disciplines scolaires*. Ils avaient été 38% à répondre qu'il s'agissait d'un travail relevant d'une seule discipline, 38% également d'un travail en dehors des disciplines scolaires, et 24% seulement d'un travail à cheval entre plusieurs disciplines. La composition de l'échantillon ne semble avoir joué aucun rôle dans ce cas, les résultats de la présente étude ne changeant guère après correction par la région linguistique et le sexe. La façon dont a été posée la question semble toutefois plus importante, et entrave la comparaison entre les deux études sur ce point. Il est probable que bien des élèves se sont sentis davantage poussés par la liste qui leur était proposée à indiquer un rattachement que dans l'étude EVAMAR I. Ce qui pourrait expliquer que 7,9% seulement d'entre eux n'ont pas relié leur travail à une matière.

Outre la question de savoir si le travail de maturité a été monodisciplinaire ou pluridisciplinaire, nous nous sommes aussi intéressés au domaine dont les élèves font relever leurs travaux. Il convient pour cela d'analyser plus précisément les indications de matière données à ce propos. Le tableau VI.11 montre comment les élèves situent leurs travaux parmi les grands groupes de disciplines⁶⁰. L'allemand, le français, l'italien, l'anglais, l'espagnol, le russe, le latin et le grec sont ici réunis dans la catégorie des langues premières et secondes; les mathématiques, les mathématiques appliquées, la physique, la chimie et la biologie ainsi que l'informatique dans les sciences de la nature; l'histoire, la géographie, la philosophie, l'économie, les sciences politiques, le

⁵⁹ La comparaison n'est pas directement possible au sens strict entre les résultats des deux études, car la composition des échantillons et le moment de l'enquête dans l'année scolaire diffèrent sensiblement. De plus, les données d'EVAMAR I étaient toutes corrigées en fonction de la région linguistique et du sexe, alors que nos résultats s'appuient sur des données en majeure partie non corrigées. Même après correction par la région linguistique et le sexe, les écarts persistent entre nos résultats et ceux de l'étude EVAMAR I.

⁶⁰ Pour les travaux interdisciplinaires, certains élèves peuvent apparaître plusieurs fois; d'autres, qui n'ont pas proposé de classification, ne pas apparaître du tout.

droit, la pédagogie, la psychologie, la sociologie et la religion dans les sciences humaines et sociales; les arts appliqués, l'histoire de l'art, le théâtre, la musique et les sports dans les arts et l'éducation physique⁶¹.

Tableau VI.11: Choix du sujet selon la discipline générale (en pour-cent, réponses multiples admises)

Groupe de disciplines	Pourcentage des travaux de maturité
Langue première et seconde	30,9
Sciences de la nature	21,8
Sciences humaines et sociales	31,5
Arts et éducation physique	15,3

Sur les 802 liens indiqués par les auteurs de travaux de maturité avec un domaine, 86,9% se référaient précisément à l'un des quatre groupes de disciplines. Au sein de chaque groupe, les réponses mentionnaient parfois plusieurs disciplines, surtout en sciences humaines et sociales: un peu plus d'un cinquième de ces travaux de maturité serait à rattacher à plus d'une discipline du groupe, selon leurs auteurs. En sciences de la nature, cette proportion n'est que de 15,5%, voire de 5,5% pour les arts et la musique, et de 3,2% seulement pour les langues. Cela tient certainement en partie au nombre variable de disciplines et aux diverses combinaisons possibles au sein de chaque groupe; mais dans le groupe des langues, on distingue une nette tendance à se cantonner dans une discipline. Un certain nombre de réponses transcendent aussi les groupes de disciplines; dans 13,1% des cas, les réponses des élèves empiètent sur plusieurs groupes de disciplines, ce qui permet de conclure avec une bonne certitude au caractère interdisciplinaire du travail.

Au cours de leur scolarité, les élèves ont surtout travaillé dans un cadre compartimenté en matières bien distinctes. C'est pourquoi on ne saurait s'attendre à une forte proportion de travaux relevant de plusieurs disciplines. On est pourtant frappé de constater à quel point il est rare qu'un travail soit qualifié d'interdisciplinaire. Cela pourrait s'expliquer par les conditions de réalisation, la dimension interdisciplinaire n'étant explicitement requise que dans les écoles romandes (cf. section 3).

En Suisse romande, la proportion de travaux monodisciplinaires est la même qu'ailleurs, mais celle des travaux interdisciplinaires et un peu plus élevée (35,1%), et celle des travaux ne relevant d'aucune discipline plus faible. Cela pourrait provenir du fait que les élèves romands ont plus explicitement envisagé l'option de l'interdisciplinarité, compte tenu des conditions dans lesquelles ils étaient placés, et ont donc plus volontiers rattaché leur travail à plusieurs matières.

4.3 Choix du sujet du travail de maturité

Trois quarts environ des élèves interrogés indiquent avoir librement choisi le sujet de leur travail de maturité, un peu moins de 14% avoir dû l'insérer dans un cadre thématique prédéterminé, et 9% avoir dû le choisir dans une liste. Si l'on fait le rapprochement avec le cadre fixé par les écoles, il apparaît que les élèves pourraient se considérer comme libres même quand ils choisissent un sujet lié à un thème ou à une matière scolaire; 35% seulement des écoles admettent le libre choix du sujet, 47% circonscrivent le choix dans un thème ou une matière, et 6% seulement donnent une liste de sujets possibles.

⁶¹ Ce rattachement indiqué par les élèves ne coïncide pas nécessairement avec les étiquettes données aux matières dans les programmes. La sociologie, par exemple, n'est pas une matière scolaire.

D'une région linguistique à l'autre, les élèves évaluent très diversement leur liberté; c'est en Suisse alémanique qu'elle domine. En Suisse romande, le choix se fait plus fréquemment dans un cadre thématique ou à partir d'une liste. Et au Tessin, les deux formules sont à peu près aussi répandues l'une que l'autre (tableau VI.12).

Tableau VI.12: Choix du sujet du travail de maturité selon la région linguistique (en pour-cent)

Mode de choix	CH alémanique	CH romande	Tessin	Total
Liberté de choix	94,8	45,9	43,6	75,9
Cadre thématique	3	29,1	40	13,7
Liste de sujets	1,1	23,3	14,5	9,1
Autres	1,1	1,7	1,8	1,3

En fait, au vu des conditions qu'ils fixent, tous les gymnases alémaniques admettent le libre choix du sujet. Un quart seulement d'entre eux prévoient en outre la possibilité de choisir sur des listes. Mais ce qui frappe, ce sont les écarts d'évaluation des élèves constatés par rapport aux gymnases tessinois, dans lesquels le travail de maturité est organisé en séminaires; les élèves peuvent suivre un séminaire sur l'un des sujets généraux offerts, et choisir dans ce cadre le sujet de leur travail de maturité – ce qui semble limiter leur choix. Mais ce n'est pas ainsi qu'ils voient les choses, puisque 43,6% d'entre eux ont indiqué avoir librement choisi leur sujet. En Suisse romande, les conditions varient d'un canton à l'autre: le sujet est parfois librement choisi, parfois sélectionné sur une liste, parfois circonscrit dans un cadre thématique.

Cela explique peut-être aussi les écarts constatés avec la première étude EVAMAR, selon laquelle 8% des élèves tessinois seulement avaient librement choisi leur sujet, et 65% des élèves romands avaient dû choisir un sujet sur une liste ou le définir dans un cadre thématique. On dirait donc que les modalités de choix ont changé.

Dans neuf cas sur dix, les élèves choisissent un sujet parce qu'il les intéresse. Ce qui n'est pas très étonnant, car on imagine mal des élèves travaillant de façon autonome pendant une période prolongée sur un sujet qui ne les intéresse guère. Quelque 30% des élèves interrogés indiquent qu'ils ont tenu compte de leurs connaissances préalables; un quart d'entre eux ont voulu relever un défi. Un peu moins de 20% ont évoqué leurs projets d'études supérieures. Un élève sur dix seulement aurait été conduit à choisir son sujet par la possibilité de faire équipe avec certains camarades. Les conseils des parents ou de proches ont aussi eu une influence modérée, tout comme l'affinité avec un enseignant. La facilité prévisible du projet ne semble pratiquement jamais avoir été un critère: 3,2% des élèves interrogés indiquent avoir songé à un sujet leur facilitant autant que possible la tâche.

Tableau VI.13: Motifs de choix du sujet (en pour-cent, réponses multiples admises)

Motif	Pourcentage des élèves interrogés
Intérêt personnel	92,4
Connaissances préalables	29,8
Défi	24,5
Etudes supérieures envisagées	18,9
Equipe	9,9
Conseils enseignants	6,7
Conseils parents	6,4
Conseils camarades	6,2
Personne suivant le travail	6,1

Les résultats de l'étude EVAMAR I étaient proches des nôtres: 90% des élèves interrogés alors avaient évoqué l'intérêt que leur inspirait le sujet, 34% leur connaissance préalable du domaine. Et la proportion de ceux qui s'étaient dits attirés par la facilité du travail, comme celle de ceux qui avaient suivi les conseils de parents, d'enseignants ou de camarades, étaient inférieures à 10% (Pagnossin, Aliata & Dozio, 2005 p. 185).

C'est donc surtout l'intérêt ressenti pour le sujet qui détermine son choix, ce qui n'a rien de bien étonnant. Même lorsque les conditions qui leur sont offertes les restreignent, les élèves ont souvent l'impression qu'ils ont choisi selon leurs préférences. On peut donc interpréter l'attrait de la connaissance préalable du domaine ou du défi à relever comme une composante additionnelle de l'intérêt. La direction que prévoit de prendre l'élève dans ses études supérieures semble être un critère fonctionnel; mais le calcul utilitariste ne conduit pas pour autant à choisir un sujet parce qu'il sera plus facile à traiter. Toutes ces indications montrent que les élèves éprouvent en majorité un authentique intérêt à la perspective d'explorer sérieusement le sujet de leur projet.

4.4 Evaluation du travail de maturité

Quelque 86% des élèves consultés ont obtenu à leur travail de maturité une bonne ou une très bonne note, 12,5% une mention «suffisant», et 1,3% seulement une mention «insuffisant». Cela dit, les critères d'évaluation changent d'un canton, voire d'une école à l'autre: la note ne s'applique souvent pas qu'au produit écrit, elle englobe également la présentation orale des résultats. Le coefficient des composantes orale et écrite dans la note varie aussi; les conséquences d'une note insuffisante sont propres à chaque établissement (cf. section 3).

Tableau VI.14: Notes données aux travaux de maturité (en pour-cent)

Note	% des travaux de maturité
sehr gut, très bien, notes 6 et 5,5	46,5
gut, bien, notes 5 et 4,5	39,7
genügend, suffisant, note 4	12,5
ungenügend, insuffisant, note inférieure à 4	1,3

Comme on peut s'y attendre, compte tenu des bonnes notes données par la plupart des enseignants, les élèves sont en majorité d'accord avec l'évaluation de leur travail. Plus de la moitié (52,4%) estiment que leur note est parfaitement justifiée, un tiers (32,7%) relativement justifiée; 11,1% jugent que leur note n'est pas vraiment justifiée, et 37% qu'elle est injustifiée. On observe bien sûr un parallélisme très net entre la note elle-même et le degré de satisfaction qu'elle inspire. Le coefficient de corrélation de 0,52 (Kendall tau) est tout à fait significatif à cet égard.

Sur les 960 personnes qui ont fait savoir si elles avaient une claire connaissance des critères d'évaluation avant de se mettre au travail, 82,9% ont répondu par l'affirmative. Quant aux 17,1% d'élèves qui ont répondu par la négative, on ne sait pas si c'est parce qu'ils ne s'étaient pas informés convenablement, pour une raison ou pour une autre, ou parce que les critères ne leur avaient pas été communiqués. A en juger par les cadres prévus par les écoles, il y a diverses façons de fixer les critères d'évaluation: deux tiers des établissements les communiquent par écrit, un tiers à peine s'en remettent aux enseignants assurant le suivi pour définir les critères et les négocier avec les élèves. Deux écoles (7%) ne donnent pas d'indications sur l'appréciation et la notation. Il est toutefois très net que parmi les élèves qui ont déclaré ne pas avoir eu connaissance des critères, plus des trois quarts auraient bien aimé les connaître, et 22% seulement trouvaient la chose inutile. Parmi les

élèves qui avaient disposé par avance des critères d'évaluation, la grande majorité jugeait cela utile (54,5%) ou très utile (39,6%), et 5,9% seulement n'en voyaient pas l'utilité.

4.5 Suivi des travaux

La plupart des élèves interrogés (59,1%) ont pu choisir eux-mêmes la personne qui encadrerait leur travail de maturité, en fonction du sujet ou de la discipline; mais pour 28% d'entre eux, c'est l'école qui l'a choisie; et dans 20% des cas, la sympathie personnelle a joué un rôle (réponses multiples admises). On note que la personne assurant le suivi du travail est nettement plus souvent nommée par l'école en Suisse romande, est plus fréquemment choisie en Suisse alémanique et au Tessin en fonction du sujet ou de la discipline. Les directives institutionnelles du Tessin conduisent à penser toutefois que le choix d'un projet est prédéterminé par celui du sujet général d'un séminaire. Il s'agit donc de la personne couvrant le sujet général dans le cadre de l'organisation en séminaires; le choix de la personne encadrant le travail de maturité passe ainsi par le choix du sujet, mais il est limité par le séminaire qu'offre l'enseignant concerné. Ainsi s'explique la forte proportion d'élèves tessinois qui indiquent avoir choisi la personne qui suivrait leur travail sur la base du sujet ou de la discipline.

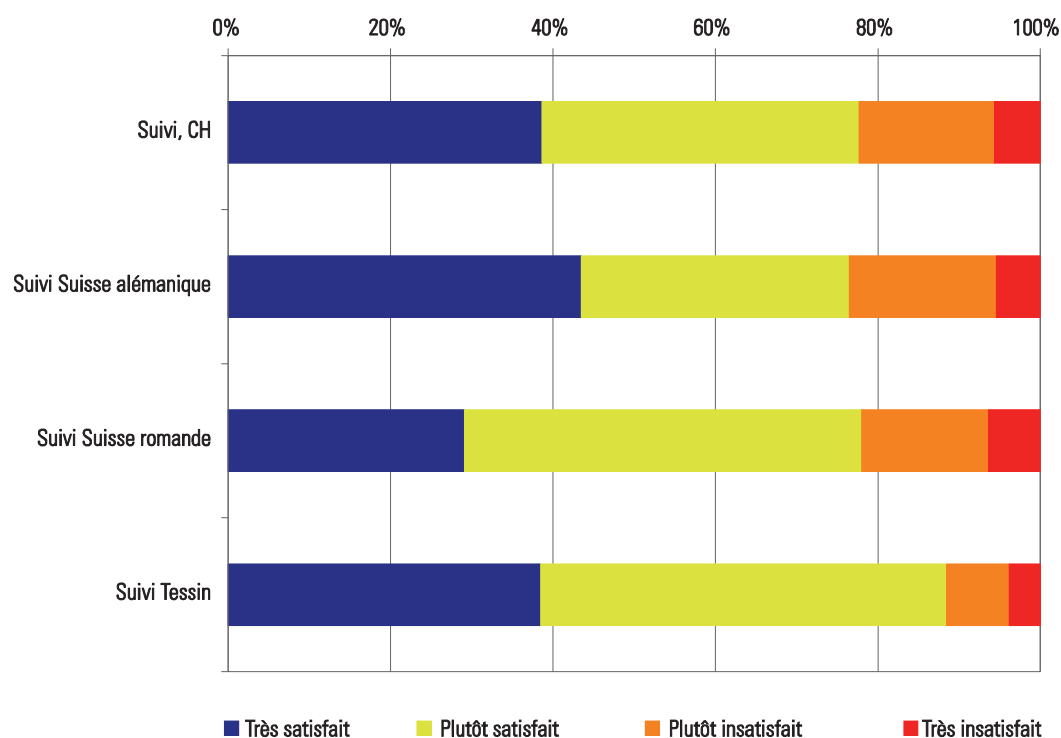
Tableau VI.15: Choix de la personne encadrant le travail de maturité selon la région linguistique (en pour-cent, réponses multiples admises)

Mode de choix du sujet	CH alémanique	CH romande	Tessin	Total
Imposé	11,2	64,1	7,3	28,2
Selon le sujet ou la discipline	73,4	27,2	85,5	59,1
Sympathie personnelle	26,5	11,1	1,8	20,0

Le suivi du travail de maturité recourt à diverses formes de communication. Plus de 90% des élèves mentionnent des entretiens individuels, un quart d'entre eux ont pris part à des entretiens de groupe, et la moitié ont été en contact par courrier électronique avec l'enseignant concerné.

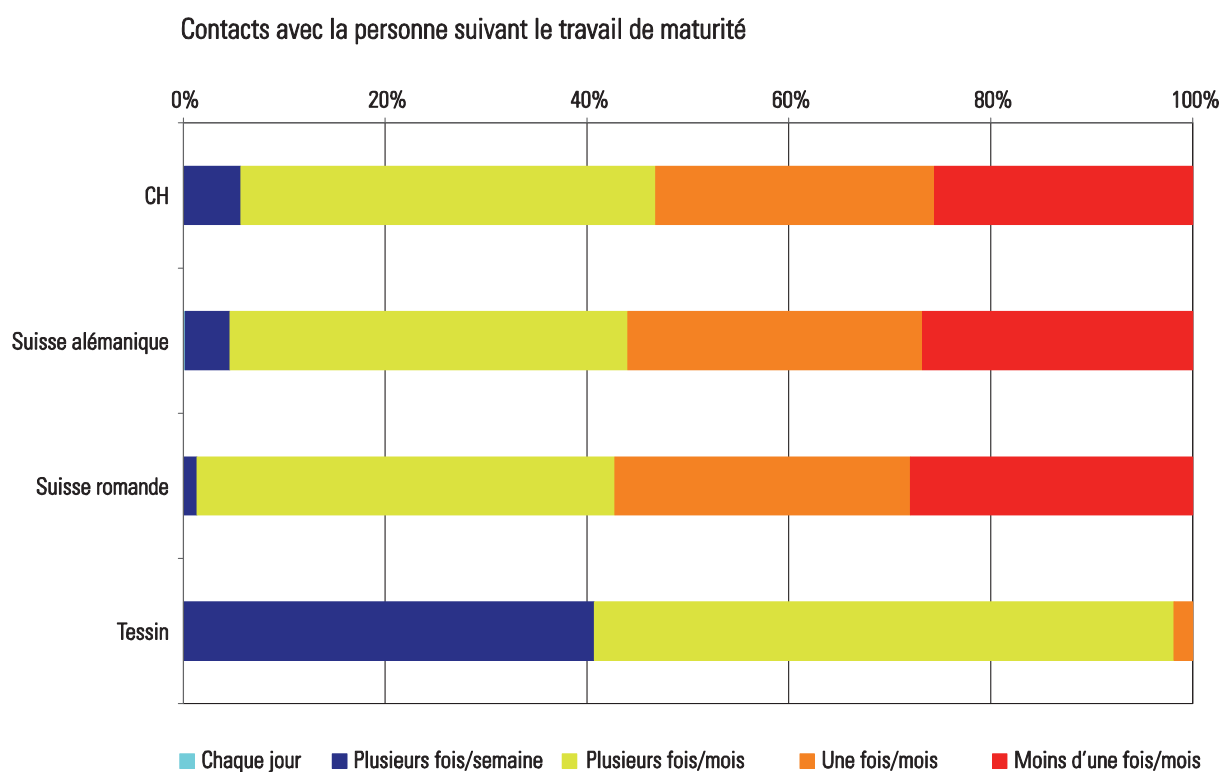
Dans l'ensemble, les élèves sont satisfaits de l'appui qu'ils ont reçu de l'enseignant ou des enseignants par lesquels ils étaient suivis. La figure VI.9 présente les appréciations selon la région linguistique. Les élèves des deux écoles tessinoises en particulier se sont sentis très bien encadrés; mais en Suisse romande aussi, malgré une appréciation relativement plus critique, la grande majorité des élèves interrogés ont été contents ou très contents de l'appui qu'ils ont reçu. L'enquête EVAMAR I avait donné des résultats similaires, avec deux tiers des élèves interrogés se disant satisfaits de l'accompagnement fourni par leurs enseignants (Pagnossin et al., 2005, p. 187).

Figure VI.9: Satisfaction éprouvée à l'égard de l'encadrement fourni par les enseignants (en pour-cent)



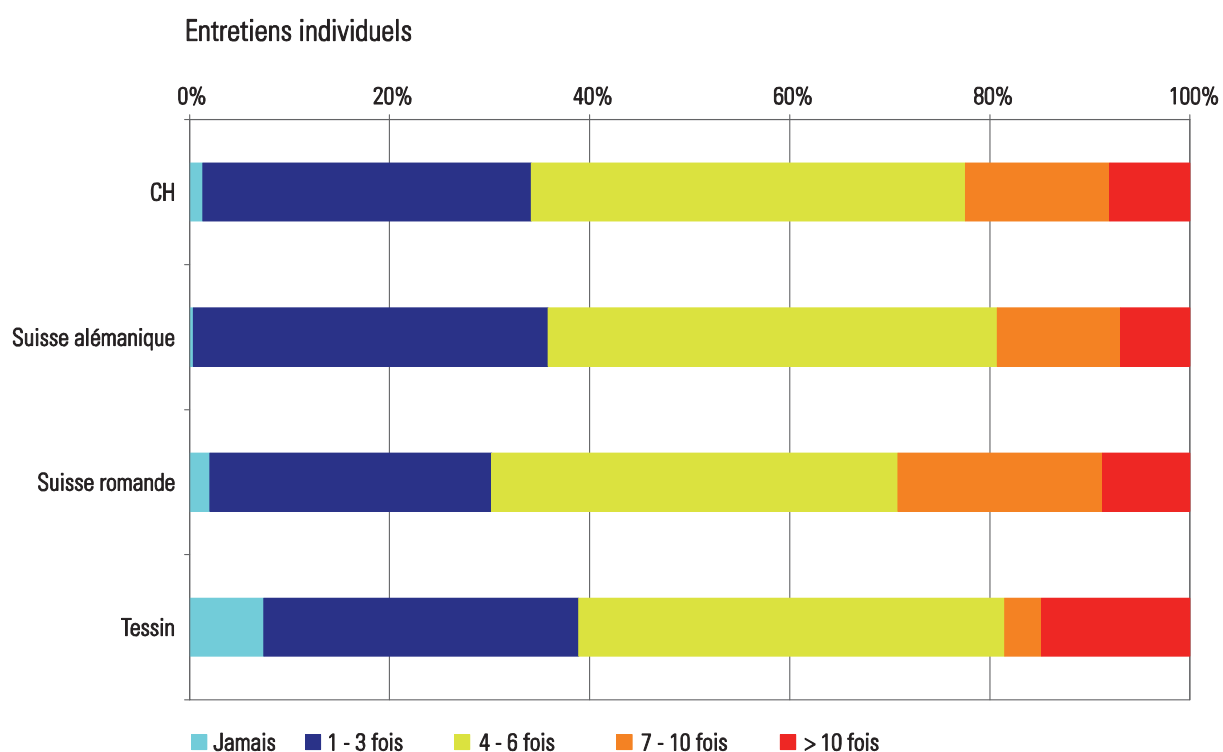
Une grande partie des élèves ont eu plusieurs fois par mois des contacts avec l'enseignant qui les suivait, un quart d'entre eux en moyenne un contact par mois, et encore autant moins d'un contact par mois. La comparaison régionale (figure VI.10) fait bien ressortir l'intensité du suivi dans les écoles tessinoises, ce qui devient moins étonnant si l'on se souvient que le travail de maturité s'intègre dans un séminaire au Tessin. Dans les deux autres régions linguistiques, les résultats sont très similaires; d'un côté comme de l'autre, on note que la fréquence des contacts pédagogiques est bien souvent relativement faible en cours de projet. Nos données ne nous permettent pas de déterminer si le phénomène s'explique par le faible nombre des contacts offerts par les enseignants ou par l'absence de besoin chez les élèves. Mais nous reviendrons sur ce point à propos de la satisfaction éprouvée par les élèves à l'égard de leur encadrement.

Figure VI.10: Fréquence des contacts avec la personne suivant le travail de maturité (en pour-cent)



On obtient un tableau plus précis en examinant la fréquence des entretiens individuels (figure VI.11). Un tiers des élèves interrogés indiquent qu'ils ont eu de un à trois entretiens en cours de travail, 43% en ont eu de quatre à six. Dans plus d'un cinquième des cas, le suivi a été fréquent ou très fréquent.

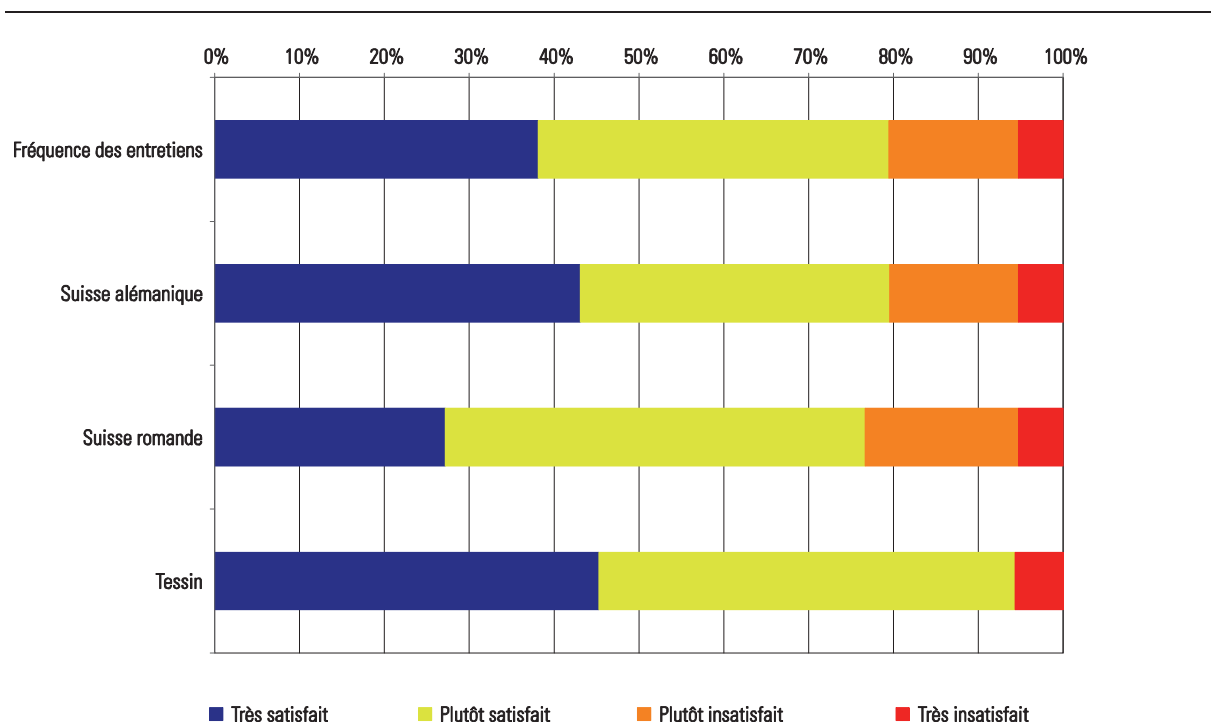
Figure VI.11: Fréquence des entretiens individuels de l'élève avec la personne suivant son travail de maturité (en pour-cent)



Bien des élèves ont donc eu peu d'entretiens individuels (figure VI.12). Comme nous l'avons dit plus haut, la fréquence des contacts ou des entretiens individuels ne permet pas de déterminer si l'offre est suffisante ou si elle répond bien à une demande. Il est alors indispensable aussi de se demander si les élèves en sont satisfaits.

La figure VI.12 présente la distribution de la satisfaction éprouvée à l'égard de la fréquence des entretiens. Au total, 80% environ des élèves interrogés s'en déclarent satisfaits ou très satisfaits. Au Tessin, la proportion est encore plus élevée, sans doute là encore en raison de l'organisation en séminaires. On observe une correspondance très marquée entre la fréquence des contacts et des entretiens individuels d'une part, et la satisfaction qu'elle inspire de l'autre. L'étude EVAMAR I avait donné un résultat proche, mais un peu inférieur: 71% des élèves s'étaient en effet déclarés satisfaits à cet égard (Pagnossin et al., 2005, p. 187).

Figure VI.12: Satisfaction des élèves à l'égard de la fréquence des entretiens (en pour-cent)



En ce qui concerne la nature des conseils obtenus (tableau VI.16), la moitié environ des élèves ont indiqué avoir fait appel à l'enseignant sur le choix du sujet, autant encore sur des questions de fond, et autant aussi sur des problèmes de structure et d'articulation du travail; 37% ont obtenu de l'assistance sur des questions de forme, et un quart ont demandé des renseignements sur des sources d'information et des méthodes de travail. Un cinquième ont eu besoin d'aide pour gérer leur temps et rédiger leur rapport. Les problèmes de motivation ont été moins souvent évoqués (12%). L'utilisation de l'informatique et de l'Internet ne semble pas non plus avoir nécessité une aide bien fréquente (6%).

Tableau VI.16: Aide demandée à l'enseignant (réponses multiples admises)

Nature de l'aide demandée	% d'élèves
Choix du sujet	16,7
Formulation de la question	49,9
Gestion du temps et organisation du travail	21,5
Informatique et Internet	6,0
Contenu du travail (aide sur le fond)	53,3
Sources d'information	27,4
Méthodes de travail	24,7
Rédaction du travail	20,1
Structure et articulation du travail	51
Motivation, crises	12,1
Questions de forme (notes de bas de page, bibliographie, citations, etc.)	37,1
Divers	4,4

L'enseignant a donc dû donner des conseils de toutes sortes. L'incertitude et le besoin de conseils ne se concentrent pas sur une phase clairement délimitée. On peut toutefois distinguer trois temps: les conseils initiaux visant à préciser les questions; l'élucidation de problèmes de fond en cours de travail; et enfin les questions de structuration et de présentation du texte en phase de rédaction. L'assistance n'est moins souvent demandée que sur les sources électroniques, en cas de crise de motivation et sur le choix du sujet. En ce qui concerne les sources électroniques, cela n'est guère étonnant, sachant que les élèves les maîtrisent souvent mieux que les enseignants. Le phénomène s'explique bien aussi pour ce qui est de la motivation et du choix du sujet quand on se rappelle que la plupart des élèves disent avoir surtout choisi leur sujet par intérêt personnel – ce qui veut aussi dire que la motivation ne serait pas vraiment un problème.

L'examen de la nature de l'assistance fournie n'explique pas pour autant les raisons pour lesquelles il y est recouru ou non, et s'il faut imputer la chose à la qualité des conseils ou aux besoins des élèves. C'est pourquoi nous avons demandé aux élèves dans le questionnaire s'ils avaient trouvé cette assistance utile. Les réponses sont presque toutes au-dessous de la valeur médiane de l'échelle (2,5), et on peut donc les considérer comme positives. La satisfaction est moindre (2,77) uniquement pour ce qui est de l'assistance en *informatique et sur l'Internet* – ce qui n'a rien de très surprenant, les élèves connaissant d'habitude bien les logiciels usuels, comme on l'a vu. L'aide reçue à la préparation du travail écrit et en cas de crise de motivation est perçue comme relativement moins utile. On peut penser que les élèves éprouvent également moins de besoins dans ce domaine que dans d'autres, les élèves ayant aussi déclaré après coup avoir moins demandé d'assistance à la rédaction et en cas de crise de motivation. C'est l'aide à la formulation des questions et à la structuration du travail qui est ressentie comme la plus précieuse.

Tableau VI.17: Appréciation de l'assistance reçue (valeur moyenne, écart type)

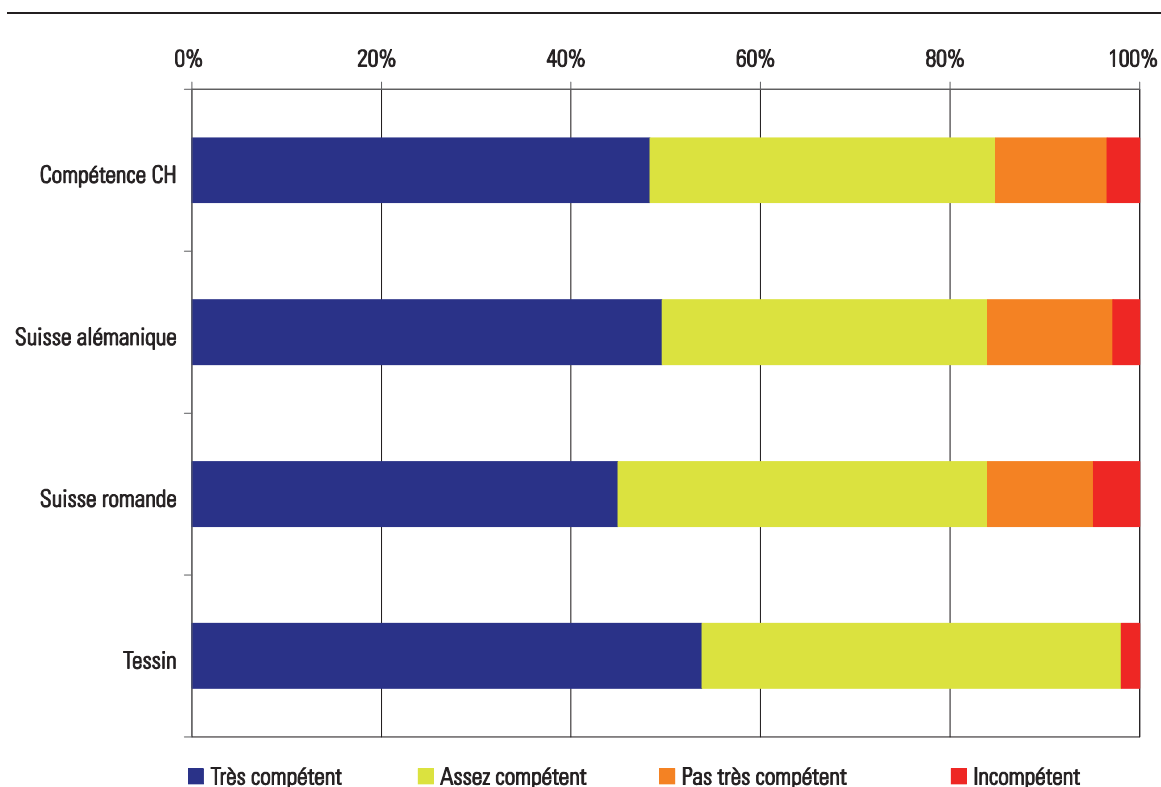
Nature de l'aide demandée	Val. moyenne	SD
Choix du sujet	2,2	1,00
Formulation de la question	1,9	0,82
Gestion du temps et organisation du travail	2,2	0,96
Informatique et Internet	2,8	1,09
Contenu du travail (aide sur le fond)	2,0	0,88
Méthodes de travail	2,2	0,93
Structure et articulation du travail	1,9	0,84
Rédaction du travail	2,4	1,04
Motivation, crises	2,3	1,09
Questions de forme (notes de bas de page, bibliographie, citations, etc.)	2,1	1,05
Divers	1,7	1,07

1= très utile, 2= assez utile, 3= pas très utile, 4= inutile

Il est important dans ce contexte de connaître l'avis des élèves sur les compétences de l'enseignant qui les accompagne, car cette impression a probablement une incidence sur la demande de soutien. De fait, on observe une corrélation positive entre l'assistance fournie sur le fond du travail et l'appréciation de la compétence de l'enseignant (Spearman's Rho = .534). De plus, les élèves qui ont demandé de l'aide pédagogique sur le contenu de leur travail ont une plus haute opinion de la compétence de l'enseignant concerné que ceux qui n'en ont pas demandé (test t bilatéral, $\alpha < .005$).

La figure VI.13 donne les évaluations en pour-cent de la compétence des enseignants sur le fond. On remarque que les élèves ont dans l'ensemble une très haute opinion de la compétence de leurs enseignants, avec les scores les plus élevés observés dans les deux écoles du Tessin.

Figure VI.13: Appréciation de la compétence de l'enseignant sur le fond (en pour-cent)



Outre l'assistance fournie par l'enseignant assurant le suivi du travail, les élèves interrogés mentionnent l'aide d'autres personnes: les parents pour un peu moins des deux tiers d'entre eux, des amis ou des connaissances pour un peu moins de la moitié, des spécialistes pour 45%. Le rôle des frères et sœurs (23%) et des camarades de classe (un peu moins de 18%) est nettement moins marqué (tableau VI.18)⁶².

Tableau VI.18: Autres sources d'aide (réponses multiples admises)

Assistance extrascolaire	% élèves interrogés
Parents	64,2
Frères et sœurs	23,0
Amis et connaissances	47,8
Camarades de classe	17,9
Spécialistes extérieurs	45,1

Les élèves qui recourent à une aide extrascolaire se tournent parfois plus souvent aussi vers l'enseignant qui les accompagne. Ceux qui vont chercher de l'aide auprès de leurs camarades en demandent moins à leurs enseignants (test t bilatéral, $\alpha < 0,05$); ceux qui consultent de surcroît un spécialiste extérieur ont des entretiens individuels plus fréquents avec leur enseignant (test t bilatéral, $\alpha < 0,05$). Il n'y a pas de corrélation observable entre les autres sources d'aide (comme les parents) et l'encadrement fourni par l'enseignant. On note à propos des contacts avec les camarades de classe ou des experts extérieurs qu'ils semblent plutôt motivés par l'intérêt

⁶² Cette présence plus faible des frères et sœurs pourrait s'expliquer par le fait que certains élèves n'en ont pas. N'ayant pas posé la question, nous ne savons pas pourquoi le rôle des frères et sœurs reste aussi effacé, par exemple par rapport à celui des parents. Les statistiques officielles montrent que 3 enfants sur 10 n'ont pas de frères et sœurs: la proportion d'élèves ayant reçu une assistance de leurs frères et sœurs remonterait alors de 32%.

éprouvé pour le fond ou la matière du travail. Il n'est donc pas étonnant que la coopération soit également plus étroite dans ce cas avec l'enseignant. Dans le cas d'un expert extérieur, il n'est pas impossible que l'enseignant ait lui-même joué un rôle dans ce contact, ce qui veut dire qu'il aurait fourni un surcroît d'assistance. Les données recueillies montrent clairement que les élèves ne considèrent pas que l'expert extérieur se substitue à leur enseignant, mais qu'il le complète dans sa spécialité. Dans cette mesure, la proportion élevée d'élèves ayant consulté un expert extérieur doit plutôt être interprétée comme un signe de bonne coopération entre élèves et enseignants.

En ce qui concerne le rôle important joué par les parents, les amis et les connaissances dans la préparation de nombreux travaux de maturité, on peut y voir le signe encourageant que les élèves auraient conscience de l'utilité de nombreux échanges avec d'autres personnes lorsqu'ils travaillent sur un projet de grande envergure. Les élèves se tournent surtout dans ce contexte vers les membres de leur famille et leurs connaissances, au-delà de leurs enseignants. Le phénomène peut faire problème dans la mesure où il introduit une ressource sociale inégalement répartie dans la population. Le capital intellectuel et matériel dont dispose chaque famille variant considérablement, cela risque de se traduire par des inégalités qu'il n'est pas sûr que les enseignants encadrant les travaux parviennent à niveler.

4.6 Cadre institutionnel et situation personnelle

Trois quarts environ (77,3%) des élèves interrogés ont confirmé que leurs écoles publient bien un guide de préparation du travail de maturité, et 93,4% qu'ils en avaient connaissance pendant leur travail. En revanche, 22,7% d'entre eux ignoraient l'existence d'un tel guide ou doutaient même expressément qu'il en existe un dans leur établissement; or ces élèves auraient aimé en avoir un, comme le montre la très forte proportion (66,9%) d'entre eux qui indiquent qu'ils auraient trouvé utile ou très utile d'en posséder un. Un résultat d'autant plus étonnant que l'analyse des cadres mis en place par les écoles indique que tous les établissements ont prévu un guide (voir section 3). D'un côté, la grande majorité des écoles semble donc s'appuyer sur le guide comme auxiliaire de préparation du travail; de l'autre, un quart environ des élèves prétendent ne pas connaître les directives de leur école. Compte tenu de leur utilité dans la préparation des travaux de maturité, de nombreuses écoles cantonales doivent donc prendre des mesures pour faire connaître et utiliser les guides. Mais ce n'est pas le cas partout: les différences sont considérables sur ce point d'un établissement à l'autre. Si cette ignorance des directives est imputable non pas aux élèves mais surtout à l'établissement, elle sera statistiquement plus répandue dans certains que dans d'autres. Effectivement, 38,6% de la variance s'explique par l'école: si tous les élèves de tel ou tel établissement indiquent avoir eu connaissance du guide, ils ne sont qu'une minorité dans tel ou tel autre.

Parmi les élèves qui connaissaient les directives de leur établissement, 61,2% les ont trouvées utiles ou très utiles, un tiers environ n'en étaient pas très satisfaits. Bien évidemment, cette appréciation de l'utilité des guides varie considérablement d'une école à l'autre, s'agissant de guides différents; 16,7% de la variance s'explique dans ce cas par l'appartenance à un établissement.

Beaucoup d'écoles organisent des séances de préparation au travail de maturité. Plus de la moitié des élèves déclarent y avoir participé; mais 42,4% rapportent que leur établissement n'en a pas organisé, même si 35,2% d'entre eux auraient trouvé cela assez ou très utile.

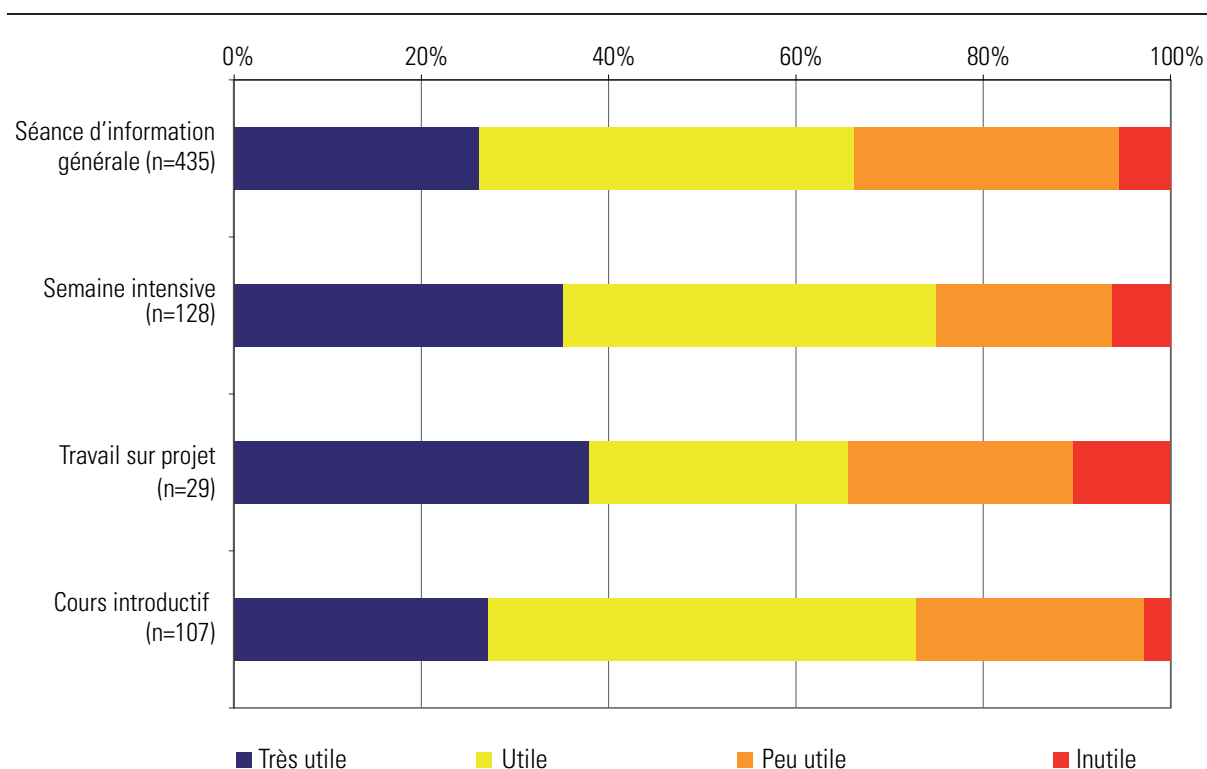
Tableau VI.19: Séances et cours préparatoires (réponses multiples admises)

Type	% élèves interrogés
Séances d'information générales ou spécifiques	82,7
Semaines intensives ou de travail sur projet	24,7
Travail en petit groupe sur projet	6,0
Cours introductif	21,7
Divers	7,0

Il ressort du tableau VI.19 qu'il s'agit surtout de séances d'information générales et spécifiques; un cinquième des élèves ont assisté à un cours introductif, un quart ont participé à une semaine intensive ou de travail sur projet⁶³. Ces formations sont en moyenne jugées assez utiles, même si un quart des élèves les trouvent dans chaque cas inutiles ou peu utiles.

On observe, comme on peut s'y attendre, des écarts entre établissements. En ce qui concerne l'appréciation portée sur les séances d'information générales ou spécifiques ainsi que sur les semaines intensives ou de travail sur projet, la variance avoisine les 17%. La figure VI.14 présente un synoptique de l'appréciation portée par les élèves sur ces cours et séances.

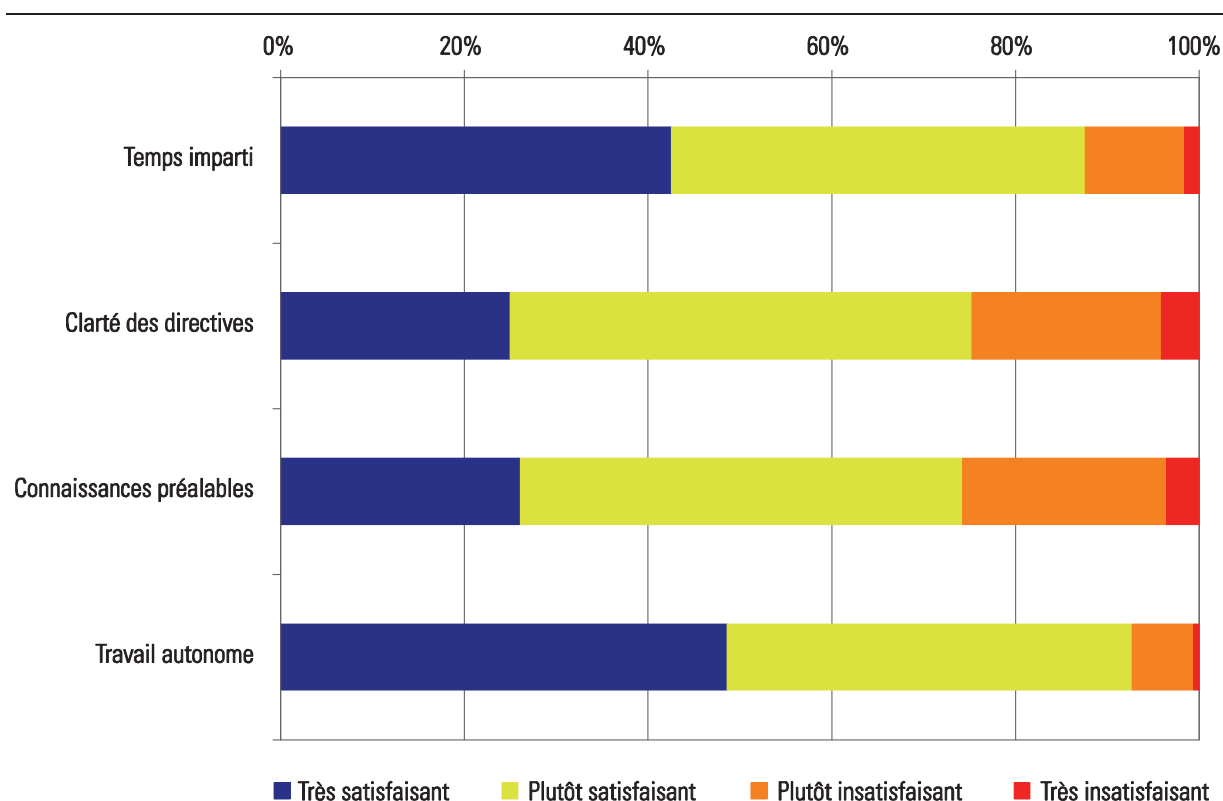
Figure VI.14: Appréciation portée par les élèves sur l'utilité des séances et cours de préparation au travail de maturité (en pour-cent)



⁶³ Il n'est pas tenu compte ici de la composante préparatoire des séances spécifiques organisées par la plupart des écoles de Suisse latine en vue de la conception et de la réalisation des travaux de maturité (voir section 3).

Notre étude des conditions de réalisation du travail de maturité nous a conduits à nous interroger sur la satisfaction qu'elles inspirent. Parmi elles figurent le temps imparti aux élèves, la clarté des directives, les connaissances préalables nécessaires à la préparation autonome d'un travail, et la capacité à travailler de façon autonome. Comme le montre clairement la figure VI.15, les avis des élèves sont en grande partie positifs. Si l'étude EVAMAR I avait donné des taux de satisfaction de 71,9% en ce qui concerne le temps imparti pour la préparation du travail de maturité, et de 55,6% pour ce qui est de la clarté des directives (Pagnossin et al., 2005, p. 187), ces proportions ont été respectivement de 87% et de 75,3% dans notre cas. Cet écart notable ne semble pas imputable à la composition des échantillons, la correction en fonction de la région linguistique et du sexe ne modifiant pas sensiblement nos résultats. Le moment de l'enquête pourrait avoir eu un effet: dans notre cas, le décalage temporel était tel que les élèves avaient déjà connaissance du résultat de leur travail de maturité, et voyaient ainsi peut-être le cadre dans lequel ils avaient travaillé sous un jour plus favorable. On peut aussi penser et espérer que les écoles ont mis à profit l'expérience qu'elles ont accumulée entre-temps sur l'organisation du travail de maturité pour améliorer le cadre de sa préparation, ce qui se refléterait dans l'appréciation des élèves.

Figure VI.15: Cadre institutionnel et individuel de préparation du travail de maturité



Les élèves sont presque tous satisfaits de leur propre aptitude au travail autonome; rares sont ceux qui répondent par «pas très bien» ou «mauvais». Ils se montrent plus critiques en ce qui concerne leurs connaissances préalables, mais il ne s'agit là encore que d'un quart environ d'entre eux. Ces résultats sont assez proches de ceux de l'étude EVAMAR I, où 68,7% étaient satisfaits de leurs connaissances préalables, et 91,1% de leur aptitude au travail autonome (Pagnossin et al., 2005, p. 187).

4.7 Appréciation du volume de travail de préparation et de l'utilité du travail de maturité

Les indications des élèves varient considérablement en ce qui concerne le temps qu'ils déclarent avoir effectivement consacré à leur travail de maturité. Un quart d'entre eux estiment y avoir passé plus de 150

heures, quelque 28% moins de 80 heures⁶⁴. La médiane se situe dans la plage 100-120 heures à 106,6. Ces écarts importants ne sauraient s'expliquer simplement par le rythme de travail de chacun; on est tenté de penser qu'ils reflètent des différences structurelles entre établissements ou entre disciplines. Dans le domaine des langues, le temps de travail a tendance à être un peu plus important (médiane = 113,7 heures), mais un peu moins en sciences sociales (médiane = 98,8). On ne constate pas de différences notables d'une région linguistique à l'autre. Pour interpréter ces résultats, il ne faut toutefois pas oublier qu'ils ne se fondent pas sur des données documentées, mais sur des volumes d'heures appréciés après coup par les élèves: il est tout à fait possible que certains d'entre eux aient surestimé et d'autres sous-estimé leur travail.

Tableau VI.20: Appréciation du volume de travail individuel effectif

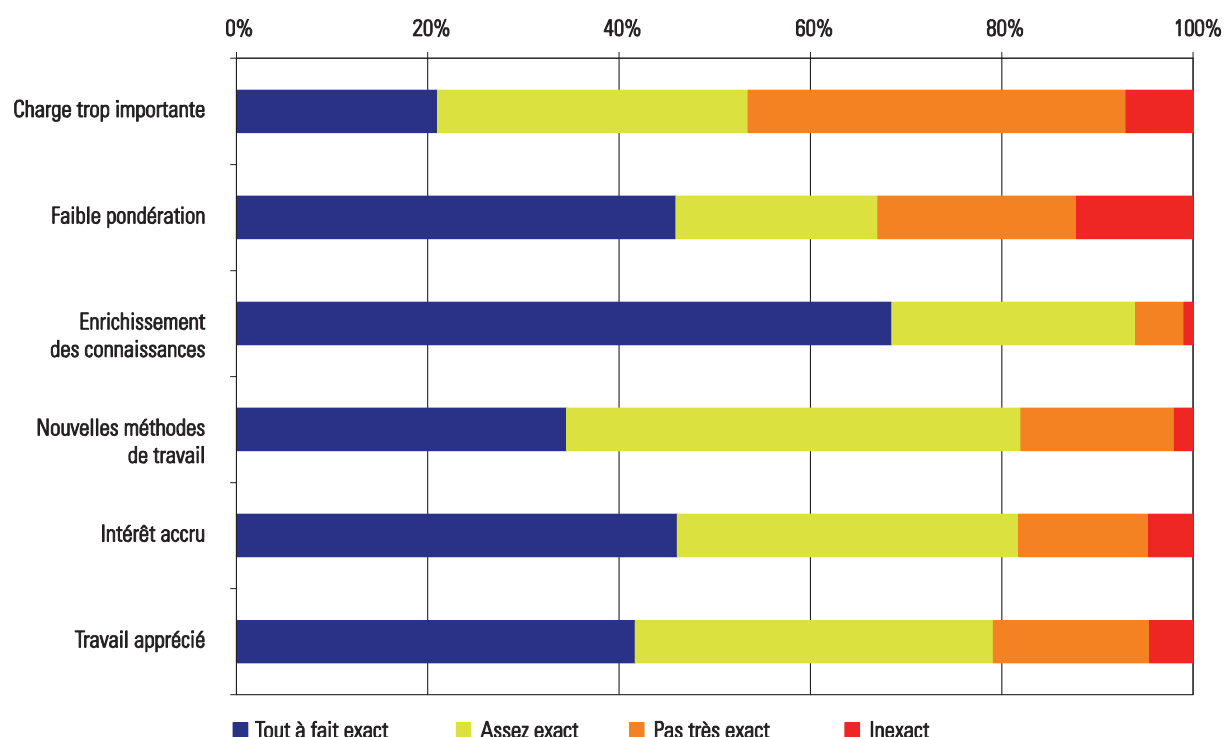
Appréciation du volume de travail (heures)	% d'élèves interrogés
Moins de 20	0,5
20-50	10,5
50-80	17,4
80-100	17,3
100-120	13,0
120-150	14,8
Plus de 150	26,5

De nombreux élèves ont pu aussi consacrer à leur travail de maturité des heures de cours normaux: 16,3 heures en moyenne. Mais là encore, la plage est très large, de 0 à 100 heures; 27,9% des élèves indiquent toutefois qu'ils n'ont jamais pu travailler à leur travail de maturité pendant les cours normaux, et 26,8% n'ont pas répondu à la question. Notre analyse des cadres fixés par les écoles confirme d'ailleurs bien que 50% environ des élèves, comme ils l'indiquent, n'auraient pas pu travailler en cours à leur travail de maturité.

De nombreux élèves trouvent cet investissement trop important, compte tenu du reste du travail qu'exige la maturité: 53 25% sont de cet avis ou tout à fait de cet avis. Ce qui peut tenir au fait que le travail de maturité ne pèse pas très lourd dans l'obtention du certificat, ce dont se plaignent d'ailleurs 67% des élèves interrogés. Outre qu'ils pensent avoir consacré un effort important à leur travail de maturité, les élèves sont en bonne part convaincus qu'ils ont beaucoup appris à cette occasion: 94% d'entre eux auraient enrichi leurs connaissances, 82% auraient appris à travailler, 81,8% ont ainsi ravivé l'intérêt qu'ils éprouvaient pour le sujet. Et 79,1% encore précisent qu'ils ont éprouvé du plaisir à faire ce travail. La figure VI.16 présente les appréciations des élèves sur le volume du travail de préparation et l'utilité du travail de maturité.

⁶⁴ En ce qui concerne l'uniformisation du temps de travail à investir, les écoles cantonales ne donnent elles-mêmes guère d'indications: rares sont celles qui le chiffrent en nombre d'heures dans leurs directives (voir section 3).

Figure VI.16: Appréciation du volume de travail de préparation et de l'utilité (en pour-cent)



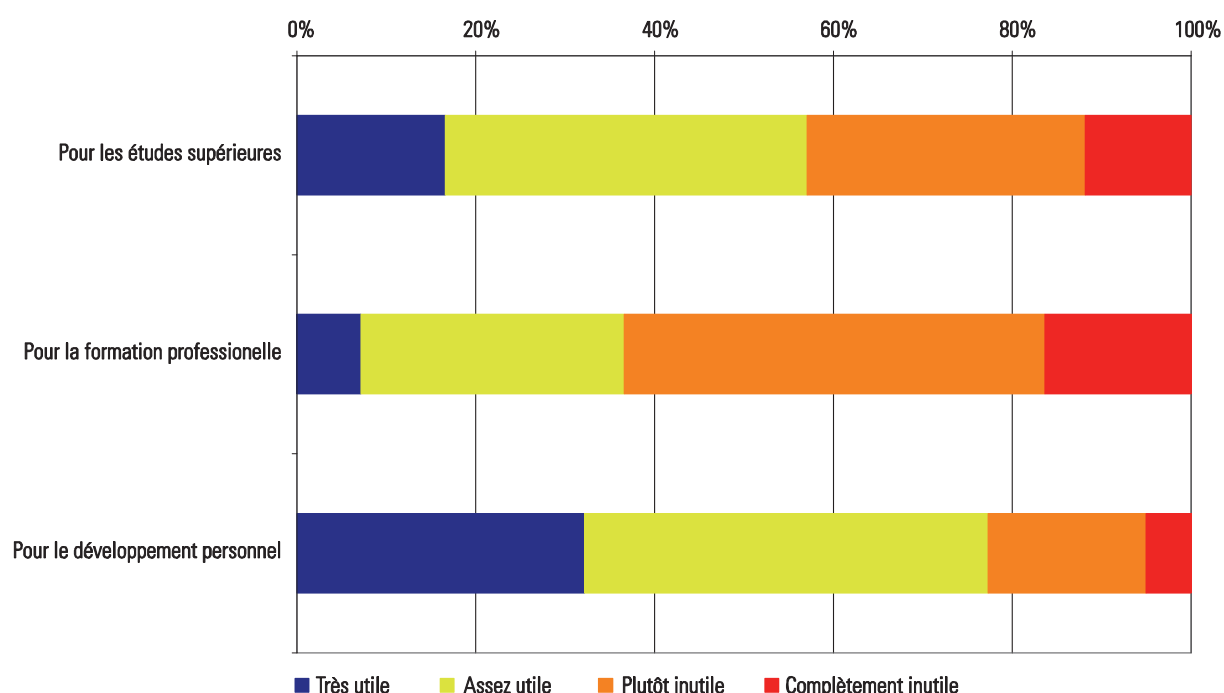
Le critère du sexe ne fait pas ressortir de différences sur le volume de travail, mais celui de la région linguistique permet de distinguer des regroupements significatifs en ce qui concerne les scores moyens notés sur trois des six aspects. Les élèves romands ne pensent pas vraiment avoir enrichi leurs connaissances; en revanche, ils trouvent que le travail de maturité a le coefficient qu'il mérite dans l'obtention du certificat. Les élèves tessinois ont dans l'ensemble éprouvé moins de plaisir à préparer leur travail de maturité que leurs camarades des autres régions linguistiques⁶⁵.

A bien des égards, ces résultats paraissent confirmer ceux de la première étude EVAMAR, qui avait montré que trois quarts des élèves rendaient un avis favorable sur le travail de maturité: 88% d'entre eux pensaient qu'ils avaient enrichi leurs connaissances, 78% qu'ils avaient travaillé dans de bonnes conditions, et 79% que leur intérêt pour le sujet avait augmenté en cours de travail. Compte tenu de l'effort qu'ils avaient dû fournir, 62% avaient trouvé que le travail de maturité ne pesait pas suffisamment lourd dans l'obtention du certificat; et la même proportion jugeait qu'il demandait un effort trop important (Pagnossin et al., 2005, p. 189). Sans qu'on puisse parler d'effets régionaux marqués, on constatait que les élèves romands se plaignaient moins que les autres du coefficient trop faible du travail de maturité dans l'obtention du certificat; et pour ce qui est du volume de travail, deux tiers environ des Alémaniques et des Romands le trouvaient excessif, alors qu'il n'y avait qu'un peu moins de 50% d'élèves tessinois de cet avis (Pagnossin et al., 2005, p. 190).

Les élèves que nous avons interrogés perçoivent une utilité relativement grande dans le travail de maturité en ce qui concerne la suite de leur itinéraire qualifiant: 57% d'entre eux le jugent utile ou très utile. Cette utilité paraît plus modeste pour la formation professionnelle, avec 38% environ d'élèves jugeant le travail de maturité utile ou très utile. Mais c'est surtout au niveau du développement personnel qu'il est ressenti comme bénéfique, avec plus de trois quarts des élèves rendant un avis favorable ou très favorable (figure VI.17).

⁶⁵ Il n'est pas impossible que le choix des mots ait joué un rôle ici: le questionnaire allemand parlait de *Spass* (qui implique la notion de plaisir ou d'amusement), le français utilisait le terme d'«appréciation», et l'italien suggérait le divertissement (*divertimento*).

Figure VI.17: Appréciation de l'utilité du travail de maturité dans l'itinéraire qualifiant et le développement personnel



Pour se faire une idée plus précise des apports du travail de maturité, on a demandé aux élèves d'évaluer ses effets sur certaines de leurs compétences. Au vu des valeurs moyennes du tableau VI.21, on s'aperçoit que le travail de maturité n'aurait pas d'effets négatifs aux yeux des élèves – même si cela ne veut pas encore dire qu'il aurait atteint ses objectifs pédagogiques. Certaines compétences auraient indéniablement été moins développées au cours du projet, aux dires des élèves, notamment le travail d'équipe, la capacité d'abstraction, la réflexion logique, l'imagination et la pensée en réseau. Il va de soi que les travaux en solitaire encouragent moins le travail d'équipe que les travaux de groupe; la moyenne est de 1,8 sur les 54 personnes qui avaient réalisé un travail en équipe, mais de 2,7 chez les autres. Les scores varient aussi en fonction du sujet pour ce qui est de la capacité d'abstraction, de la réflexion logique, de l'imagination et de la pensée en réseau; mais les résultats ne confirment pas l'hypothèse selon laquelle ce serait surtout le cas en mathématiques et dans les sciences de la nature, et la distribution ne change guère si l'on se limite aux travaux de maturité relevant des sciences de la nature en branche principale.

Tableau VI.21: Apports du travail de maturité

Que vous a apporté le travail de maturité en ce qui concerne les compétences suivantes?	Moyenne	SD
Travailler en équipe	2,7	0,84
Se discipliner	1,9	0,72
Effectuer une recherche par soi-même	1,8	0,69
Porter un jugement personnel	1,9	0,70
Communication écrite	2,1	0,75
Communication orale	2,2	0,77
Imagination	2,4	0,76
Réflexion logique	2,5	0,73
Capacité d'abstraction	2,6	0,68
Penser en réseau	2,3	0,76
Maturité personnelle	1,9	0,77

1= très bénéfique, 2= assez bénéfique, 3= apport nul, 4= assez négatif, 5= très négatif

Pour présenter plus clairement les avis des élèves sur les apports du travail de maturité et poursuivre leur analyse, nous avons préparé sur chaque compétence des échelles faisant ressortir les effets qu'aurait eus sur elle le travail de maturité. Nous avons procédé à cet effet à une analyse factorielle exploratoire qui a montré qu'il était possible de condenser ces compétences sur trois variables fondamentales. La première a trait à l'organisation autonome du travail: recherche indépendante, autodiscipline, jugement personnel, communication écrite, maturité personnelle. La seconde est centrée sur les compétences cognitives: capacité d'abstraction, réflexion logique, imagination et pensée en réseau. Et la troisième englobe les compétences sociales, formées seulement de la capacité à travailler en équipe et de la communication orale. Ces trois variables permettent de rendre compte de 57,7% de la variance, avec 23,2% pour la première, 22,7% pour la deuxième et 12,2% pour la troisième après première rotation Varimax.

Les variables des deux premiers facteurs ont été réunies dans des échelles de totalisation, afin de les rendre utilisables dans d'autres analyses. Ces deux échelles acquièrent ainsi une cohérence interne suffisante, avec un coefficient de fiabilité interne de Cronbach (alpha) de 0,75. Nous n'avons pas jugé bon de créer d'échelle pour la *compétence sociale*, car elle ne contient que deux items.

Tableau VI.22: Echelles «organisation du travail» et «compétences cognitives»

«Quels ont été les effets de la préparation de votre travail de maturité sur les aspects suivants de vos compétences?»

Item		Valeurs	Réponses possibles
Organisation du travail	Recherche autonome	Cas valables: n=902 Cronbach's Alpha: .755 M=1,9 SD=.51	Très positifs Assez positifs Nuls Assez négatifs Très négatifs
	Autodiscipline		
	Jugement personnel		
	Communication écrite		
	Maturité personnelle		
Item		Valeurs	
Compétences cognitives	Capacité d'abstraction	Cas valables: n=894 Cronbach's Alpha: .751 M=2,49 SD=.56	
	Réflexion logique		
	Imagination		
	Penser en réseau		

Sur l'échelle de *l'organisation autonome du travail*, la valeur moyenne totale de 1,9 (SD= 0,51) est nettement positive. Les élèves ont donc en moyenne l'impression d'avoir bénéficié d'un apport utile à ce niveau. La moyenne de 2,4 (SD = 0,56) enregistrée sur l'échelle des *compétences cognitives* est par comparaison moins positive. La comparaison régionale ne fait pas ressortir de différences. Seuls les élèves du Tessin pensent avoir bénéficié un peu plus de leur travail en termes de compétences cognitives. La discipline dont relève le travail de maturité ne semble pas avoir joué de rôle. On notera toutefois la corrélation négative dans les deux cas avec le temps investi: plus les élèves se souviennent avoir consacré de temps à leur travail de maturité, moins ils estiment que ce dernier leur a apporté au niveau de leurs compétences cognitives ou de leur capacité de s'organiser par eux-mêmes.

On remarque aussi le parallélisme étroit entre l'évaluation de l'utilité du travail de maturité dans des études supérieures ou professionnelles ultérieures ainsi que dans le développement personnel, et celle des apports obtenus par sa préparation en termes de compétences cognitives et de capacité d'organisation autonome, comme cela ressort de la synthèse du tableau VI. 23.

Tableau VI.23: Corrélation entre l'utilité et les apports perçus

Utilité pour...	Compétences cognitives	Organisation autonome
les études supérieures	0,9**	0,24**
la formation professionnelle	0,18**	0,26**
le développement personnel	0,32**	0,46**

** Corrélation significative au niveau 0,01 (bilatéral).

On dirait donc que les élèves ont conscience de la fonction de propédeutique scientifique du travail de maturité. Ceux qui pensent avoir amélioré à sa faveur leurs compétences cognitives et leur capacité de s'organiser par eux-mêmes dans leur travail perçoivent aussi en général l'utilité de ces acquis pour leur développement personnel et professionnel à venir.

4.8 Bilan partiel

L'analyse des réponses des élèves à l'enquête montre nettement que le suivi n'a peut-être pas été particulièrement intense pour la majorité des travaux, mais qu'il a été largement ressenti comme tout à fait satisfaisant. Même s'il faut tenir compte de l'effet de halo de notes en majeure partie bonnes, s'agissant d'une appréciation portée par les acteurs eux-mêmes *a posteriori*, on est en droit de penser que les établissements répondent bien dans l'ensemble aux besoins d'encadrement. La plupart des élèves interrogés se montrent également satisfaits des produits de leur travail. On remarque qu'ils sont nombreux à penser qu'ils ont enrichi leurs connaissances, et ont beaucoup appris en matière de méthodes de travail. Le plus souvent, ils se souviennent que leur intérêt s'est accru en cours de projet, et qu'ils ont éprouvé du plaisir à préparer ce travail. Il est donc naturel que la plupart des élèves jugent ensuite que le projet a contribué à leur développement personnel. En revanche, on observe un certain déficit en ce qui concerne l'enrichissement des compétences qualifiantes: l'utilité perçue est nettement inférieure en termes de compétences transférables aux études supérieures ou à la formation professionnelle.

Les élèves font preuve dans notre étude d'un taux de satisfaction légèrement supérieur sur l'ensemble du registre à celui qu'avait relevé l'étude EVAMAR I. Les écarts sont particulièrement perceptibles pour ce qui est du cadre scolaire de réalisation du travail, comme la transparence des directives ou les séances et cours supplémentaires de préparation. Ce qui peut s'expliquer par le fait que lorsque nous les avons consultés, les élèves de notre échantillon connaissaient déjà leur note, et se sont donc peut-être montrés rétrospectivement moins critiques. Mais n'oublions pas que les écoles ont accumulé de l'expérience depuis la première étude EVAMAR, ce qui pourrait compter pour beaucoup aussi; cela est d'autant plus probable qu'elles ont optimisé encore ces dernières années les conditions de préparation du travail de maturité, dans le cadre de leur processus d'apprentissage institutionnel. Mais notre étude révèle aussi qu'il reste bien souvent encore un important potentiel d'amélioration à ce niveau. Si, par exemple, il fallait atteindre une proportion supérieure de travaux interdisciplinaires, c'est un aspect sur lequel il serait souhaitable d'insister davantage dans les directives écrites comme dans l'encadrement des travaux.

5 La valeur de propédeutique scientifique du travail de maturité

Pour apprécier la valeur de propédeutique scientifique du travail de maturité, nous avons soumis chaque travail à deux experts qui en ont évalué une sélection d'aspects à la lumière de la grille de critères. Il est normal dans un tel contexte que des personnes différentes accordent plus ou moins d'importance à divers aspects pour former leur jugement (cf. Ingenkamp, 1989). Nous avons pallié le problème, sans toutefois l'éliminer, en nous appuyant sur une grille de critères clairement structurée, et par la formation soigneuse des évaluateurs. Un travail de maturité est en outre difficile à analyser pour d'autres raisons. Par exemple, il ne s'agit pas seulement de tenir compte de critères de qualité très divers, il faut aussi tenter de couvrir dans un instrument unique les traditions et conventions de diverses disciplines. Un grand nombre d'aspects importants se dérobent ainsi à la quantification exacte et compliquent la tâche des évaluateurs pour ce qui est de l'homogénéité des appréciations.

5.1 Double évaluation

Nous avons estimé que même des évaluateurs très soigneusement formés pouvaient produire des écarts d'évaluation provenant de certains caractères spécifiques des travaux de maturité examinés, impossibles à intégrer tous dans une grille de critères. On ne peut alors pas distinguer exactement, en cas de divergence, toutes les composantes de l'appréciation. En présence d'écarts importants entre les deux évaluateurs, nous n'avons donc pas tenu compte de l'appréciation relative à l'item concerné dans l'analyse du critère correspondant. Nous avons aussi été poussés à cette solution par le fait qu'un calcul fondé sur une moyenne arithmétique aurait donné une fausse impression d'homogénéité lorsque l'écart dépasse deux points de l'échelle.

Le pourcentage des critères d'évaluation des travaux de maturité concernés est d'ailleurs relativement faible, comme le montre le tableau VI. 24.

Tableau VI.24: Proportion d'écarts d'évaluation ≤ 2 points de l'échelle (pourcentage par critère)

Critères	Ecarts d'évaluation ≤ 2 (%)	Critères	Ecarts d'évaluation ≤ 2 (%)
Appréciation globale	94,1	Réflexion sur le processus de travail	83,3
Page de titre	98,2	Présentation citations	85,5
Présentation générale	97,5	Indication des sources	85,9
Tableaux et figures	90,9	Qualité des sources	90,4
Continuité du texte	91,8	Utilisation des tableaux et figures	87,7
Clarté	93,4	Problématique générale	84,6
Choix des mots	93,6	Formulation des problèmes	84,6
Expression	93,6	Insertion du sujet	86,5
Structure syntaxique	92,4	Clarté de la construction	81,2
Correction de la langue	93,1	Différenciation	86,8
Articulations formelles	94,3	Neutralité axiologique	86,8
Titre	94,0	Logique	87,3
Plan détaillé	91,3	Choix des méthodes	88,7
Volume et équilibre de l'ensemble	91,8	Description des méthodes	87,2
Hiérarchisation des idées	89,9	Conceptualisation	85,9
Cohérence de l'argumentation	90,8	Maniement des notions	90,3

La complexité du sujet a imposé de recourir à de nombreux spécialistes maîtrisant les pratiques scolaires de régions linguistiques différentes. La mesure de l'homogénéité de leurs jugements a été rendue considérablement plus difficile par le fait qu'ils travaillaient en binômes variables, et que certains n'avaient eu à examiner que peu de travaux. La corrélation intraclasse a donné une bonne valeur ICC=0,732 pour l'appréciation globale, c'est-à-dire l'impression générale que retire l'évaluateur du travail examiné. On peut aussi y lire la confirmation des pourcentages élevés d'appréciations distantes de moins de deux points sur l'échelle. En d'autres termes, on est en droit de qualifier le taux de concordance de satisfaisant.

5.2 Caractéristiques de l'échantillon

L'analyse de la valeur de propédeutique scientifique du travail de maturité se fonde sur un échantillon de 437 travaux de maturité, dont 56,8% de Suisse alémanique, 35,5% de Suisse romande et 7,8% du Tessin, avec 380 de ces travaux préparés isolément, 42 en binômes, et 15 en trinômes.

La répartition sommaire de l'échantillon selon la discipline donne le tableau suivant: deux tiers des travaux relèvent des sciences sociales (géographie et histoire, mais aussi économie, droit et politique), un peu moins de 15% des sciences humaines, selon les évaluateurs, et 17% environ des sciences de la nature et des mathématiques. Dans ce classement, les langues figurent parmi les sciences humaines.

Tableau VI.25: Répartition des travaux selon la discipline

Discipline	Pourcentage de travaux
Sciences humaines	14,8
Sciences sociales	68,1
Sciences de la nature et mathématiques	17,1

Les évaluateurs ont unanimement déclaré 55,6% de ces travaux monodisciplinaires, et 17,6% interdisciplinaires⁶⁶. Dans un quart des cas, ils n'ont pas été tous d'accord sur le rattachement d'un travail à une ou plusieurs disciplines. Un tiers des travaux ont été considérés comme surtout théoriques, 30% comme théoriques et empiriques dans une égale mesure, et un peu moins de 12% comme présentant une dominante empirique. Là encore, on est frappé du nombre de travaux sur lesquels les évaluateurs n'ont pas été d'accord: dans un quart des cas, ils divergeaient sur la proportion de théorie et de pratique. Il apparaît ainsi que la grille de critères n'était pas suffisamment différenciée sur ces distinctions, et que l'évaluation de l'équilibre entre composantes théoriques et empiriques ou monodisciplinaires et pluridisciplinaires aurait pu être plus féconde.

Pour ce qui est des soupçons de plagiat, le jugement des évaluateurs a relativement bien coïncidé avec les résultats donnés par le logiciel *Plagiarism Finder* (voir aussi 2.2.5): sur sept travaux soupçonnés de plagiat et soumis à *Plagiarism Finder*, deux n'ont donné aucune correspondance avec des textes publiés sur l'Internet, quatre ont produit entre 2% et 5% de concordance, et un seul 13% de concordance. Sur six travaux non soupçonnés mais soumis au logiciel pour vérification, ce dernier a effectivement confirmé qu'ils ne contenaient pas de passages empruntés non repérés comme tels: *Plagiarism Finder* a donné à chaque fois un taux de concordance de 0%.

Si l'on s'appuie sur la technique Google au lieu du logiciel de vérification de plagiat, on obtient une image différente, mais surtout plus différenciée. Les contrôles aléatoires supplémentaires pratiqués sur une sélection de passages soumis à Google ont révélé que tous les passages détectés par le logiciel n'étaient pas nécessairement des plagiat. Sur la vingtaine de travaux soumis à un contrôle spécial, neuf contenaient des passages non repérés par des guillemets mais récupérés tels quels sur l'Internet. Au total, sept de ces travaux avaient été soupçonnés de plagiat par l'un des évaluateurs ou les deux; dans deux cas, les évaluateurs n'avaient rien soupçonné. Onze travaux ont pu être déclarés exempts de plagiat par cette méthode; dans quinze autres, des savoirs n'appartenant pas au domaine public ont été présentés sans référence explicite à leur source au fil du texte; mais le plus souvent, la source figurait bien dans la bibliographie. A une exception près, il s'agissait donc dans tous ces travaux de défauts occasionnels de présentation des citations. Très souvent, par exemple, des réflexions reprises d'autres ouvrages n'étaient pas directement signalées, mais la source apparaissait au début d'une section ou d'un chapitre, sans claire référence à un passage précis. Fréquemment aussi, les citations n'étaient pas repérées par des guillemets. Dans la plupart des travaux, les sources Internet n'étaient données que sous forme d'URL, sans mention de l'auteur ni du titre du site concerné.

En ce qui concerne les soupçons de plagiat, il faut bien souligner que c'est l'intention de tromper qui compte: il n'y a plagiat que si la personne se pare des plumes du paon, c'est-à-dire s'attribue les connaissances ou les découvertes d'un autre par dissimulation ou en mentant sur la propriété intellectuelle de l'idée. Pour beaucoup des travaux trouvés ici en défaut, il devait s'agir d'une maladresse plutôt que d'une malhonnêteté délibérée.

⁶⁶ Il n'a pas toujours été facile de qualifier un travail de monodisciplinaire ou d'interdisciplinaire. Les travaux se cantonnent souvent surtout dans une discipline, avec quelques incursions dans d'autres. Il s'est révélé difficile en cours d'évaluation de tracer de nettes frontières entre les disciplines. La concordance n'est pas très bonne entre l'avis des évaluateurs et celui des élèves sur le nombre de matières principales et secondaires dont relèvent les travaux.

5.3 Conception des échelles

L'évaluation de la qualité des travaux de maturité se fonde sur un total de 32 items couvrant des aspects linguistiques et de fonds, ainsi que sur une appréciation globale reflétant l'impression générale de l'évaluateur. L'analyse factorielle confirme la fiabilité satisfaisante des trois échelles ci-dessous.

Tableau VI.26: Saturation factorielle des composantes de la qualité des travaux de maturité

Critère	Facteur 1	Facteur 2	Facteur 3
Page de titre			0,49
Présentation générale			0,51
Visuels (tableaux et figures)			0,82
Continuité du texte		0,81	
Clarté	0,45	0,76	
Choix des mots	0,43	0,76	
Expression		0,83	
Structure syntaxique		0,85	
Correction de la langue		0,79	
Articulations formelles	0,55		0,63
Titre	0,58		
Plan détaillé, hiérarchie des intertitres	0,70		
Longueur et équilibre de l'ensemble	0,77		
Hiérarchisation des idées	0,77	0,40	
Cohérence de l'argumentation	0,83		
Réflexion sur le travail	0,77	0,40	
Présentation des citations			0,67
Indication des sources			0,61
Qualité des sources	0,55		0,54
Fonctionnalité des tableaux et figures			0,83
Problématique générale	0,75		
Formulation des questions	0,69		0,40
Insertion du sujet	0,48	0,40	
Clarté de la construction	0,73		
Différenciation	0,63		
Neutralité axiologique	0,57		0,42
Logique	0,67	0,49	
Choix des méthodes	0,84		
Description des méthodes	0,69		
Conceptualisation	0,67		
Maniement des notions	0,57	0,59	

(Saturation factorielle >.4, rotation Varimax)

Dans le premier facteur, qui correspond à 30% environ de la variance totale, ce sont *le choix des méthodes* et *la cohérence de l'argumentation* qui arrivent en premier. Dans le deuxième, qui rend compte d'un peu moins de 20% de la variance totale, ressortent surtout *l'expression* et *la structure syntaxique*, ainsi que *la continuité du texte* et *la correction de la langue*. Et dans le troisième, qui couvre 15% environ de la variance totale, ce se sont les visuels et la fonctionnalité des tableaux et des figures qui l'emportent. Plusieurs items pèsent relativement lourd sur deux facteurs: la clarté de l'expression et le choix des mots, la hiérarchisation des idées, la réflexion sur le processus de travail, l'insertion du sujet, la logique et l'emploi de la terminologie.

Facteur I: l'échelle *qualité du contenu* regroupe 11 items couvrant des éléments théoriques et méthodologiques, ainsi que le centrage thématique et la rigueur de l'argumentation du travail de maturité. Plusieurs items en ont été exclus, tout en saturant fortement ce facteur comme d'autres. Il peut paraître bizarre au premier abord d'avoir placé la méthodologie dans le contenu; mais le choix et l'utilisation des méthodes ont des répercussions déterminantes sur le contenu d'un travail scientifique et à caractère de propédeutique scientifique.

Tableau VI.27: Echelle «qualité du fond»

Items	Valeurs	Réponses possibles
Titre	Cas valables: n=437 Cronbach's Alpha: .93 M=2,7 SD=0,92	1 Remplit pleinement les exigences 2 3 4 5 6 Ne remplit pas du tout les exigences
Plan détaillé, hiérarchie des intertitres		
Longueur, équilibre de l'ensemble		
Hiérarchisation des idées		
Cohérence de l'argumentation		
Problématique générale		
Clarté de la construction		
Différenciation		
Adéquation des méthodes		
Description des méthodes		
Conceptualisation		

Facteur II: l'échelle *qualité du fond* comporte six items couvrant l'expression et le style, mais aussi la grammaire et l'orthographe, ainsi que l'adéquation et la cohérence de la présentation des résultats à l'écrit.

Tableau VI.28: Echelle «qualité de la langue»

Items	Valeurs	Réponses possibles
Continuité du texte	Cas valables: n=432 Cronbach's Alpha: .94 M=2,5 SD=0,91	1 Remplit pleinement les exigences
Clarté		2
Choix des mots		3
Expression		4
Structure syntaxique		5
Correction de la langue		6 Ne remplit pas du tout les exigences

L'item *maniement des notions*, qui pèse relativement lourd dans ce facteur comme dans le premier, n'a pas logiquement sa place ici, compte tenu de son centrage sémantique, et ne contribue pas à la qualité de l'échelle, avec un coefficient de discrimination de 0,57.

Tableau VI.29: Echelle «qualité de la forme»

Items	Valeurs	Réponses possibles
Page de titre	Cas valables: n=437 Cronbach's Alpha: .75 M=2,5 SD=0,79	1 Remplit pleinement les exigences
Présentation générale		2
Visuels (tableaux et figures)		3
Articulations formelles		4
Présentation des citations		5
Indication des sources		6 Ne remplit pas du tout les exigences
Fonctionnalité tableaux et figures		

Facteur III: l'échelle *qualité de la forme* comporte sept items couvrant l'apparence générale du travail et la mise en forme du texte. Elle dépasse la simple impression visuelle que produit le document écrit chez son lecteur; outre la lisibilité et la clarté, elle examine le respect des règles du discours scientifique, notamment au niveau des emprunts de la propriété intellectuelle d'autrui sur lesquels s'est appuyé l'auteur du projet.

L'analyse factorielle fait ressortir des variables latentes, qui couvrent plusieurs dimensions et reflètent des aspects hétérogènes de la qualité, entre lesquels on retrouve toutefois des corrélations nettes. En fait, la corrélation positive est tout à fait significative entre les trois échelles: si l'on prend un travail présentant une bonne qualité de fond, il est très probable que sa qualité sera bonne aussi en ce qui concerne la langue et la forme.

Tableau VI.30: Corrélations entre les trois échelles (qualité du fond, de la langue et de la forme)

	Qualité du fond	Qualité de la langue	Qualité de la forme
Qualité du fond		0,67**	0,6**
Qualité de la langue	0,67*		0,53**
Qualité de la forme	0,6**	0,53**	

** La corrélation est significative au niveau 0,01 (bilatéral).

5.4 Qualité du fond

L'échelle *qualité du fond* mesure la rigueur thématique, le fondement méthodologique et la logique argumentative du travail de maturité, sans faire de distinction entre les projets théoriques et empiriques. Car la nécessité de formuler clairement les problèmes, de choisir et d'utiliser soigneusement les méthodes, et de conduire une réflexion cohérente s'applique à toutes les disciplines, et à toutes les spécialités à l'intérieur de chacune.

Dans cette échelle, la moyenne arithmétique est de $M = 2,7$ et l'écart type $SC = 0,92$. En ce qui concerne les items, on constate que le titre est en général bien trouvé, et le plan détaillé du travail réussi. La plupart des élèves savent aussi dégager nettement les points principaux, et s'en tenir à leur sujet. Les notions spécifiques sont bien définies, et utilisées en contexte. Les résultats sont un peu moins bons aux yeux des évaluateurs en ce qui concerne la différenciation de l'argumentation: les élèves ne parviennent pas toujours à emprunter plusieurs approches pour répondre à une question, ou à justifier la perspective qu'ils choisissent en arbitrant entre les différentes possibles.

En ce qui concerne le choix des méthodes et leur description, on est surpris de constater que si les évaluateurs jugent d'habitude les méthodes convenablement choisies en fonction du but et du sujet, ils se montrent plus critiques à l'endroit de leur description. Les élèves ont pris les bonnes solutions, mais ils ne les retracent pas toujours dans leur compte rendu d'une façon qui permette à quelqu'un d'autre de vérifier leur choix.

Tableau VI.31: Valeurs moyennes et écarts types des critères de l'échelle «qualité du fond»

Critères	1	2	3	4	5	6
Titre	2.1(0.97)					
Plan détaillé, hiérarchie des intertitres	2.5(0.99)					
Longueur et équilibre général	2.8(1.06)					
Hiérarchisation des idées	2.5(1.08)					
Cohérence de l'argumentation	2.8(1.13)					
Problématique générale	2.6(1.33)					
Clarté de la construction	2.9(1.61)					
Différenciation	3.1(1.35)					
Adéquation des méthodes	2.4(1.18)					
Description des méthodes	3.4(1.56)					
Conceptualisation	2.3(1.10)					

L'analyse de variance montre que la comparaison multiple ne révèle pas d'écarts significatifs de qualité du contenu entre les travaux de sciences humaines, de sciences sociales et des sciences de la nature. Le test t ne fait apparaître non plus aucune différence significative entre les sexes. Mais avec une moyenne de 2,4, les travaux d'équipe présentent une avance significative ($\alpha < 0,001$, bilatéral) sur les travaux réalisés en solitaire ($M = 2,7$). Ce qui pourrait s'expliquer par le fait qu'une équipe discute et approfondit les divers aspects du sujet avant de les intégrer dans les travaux de maturité. L'analyse de variance par région linguistique fait ressortir que les élèves romands ($M = 2,1$) ont une avance significative ($\alpha < 0,001$, bilatéral) sur leurs camarades alémaniques ($M = 3,0$); la moyenne est de $M = 2,7$ pour le Tessin, mais le nombre des cas ne permet pas de pratiquer de test de signification. Pour ce qui est des écarts entre les régions linguistiques, il est regrettable qu'il n'ait pas été possible de recruter plus de spécialistes bilingues à associer aux évaluations, car cela aurait permis de procéder







à une comparaison directe par analyse de variance entre les régions. Bien que la panoplie d'instruments soit la même et que les évaluateurs aient reçu la même formation, il n'est pas impossible que des différences culturelles entre enseignants et élèves de régions différentes se traduisent par des différences d'appréciation.

5.5 Qualité de la langue

L'échelle *qualité de la langue* comporte des items portant plutôt sur le style, et d'autres sur les façons correctes d'écrire et de manipuler la syntaxe. Il faut tenir compte ici du fait que ces deux aspects ne reçoivent pas toujours le même traitement: le respect de la grammaire et de l'orthographe est obligatoire dans toutes les sciences; mais les questions stylistiques sont bien plus importantes en sciences humaines qu'ailleurs. En sciences techniques, en sciences de la nature, en mathématiques, il peut être nécessaire de sacrifier le raffinement stylistique à la précision, notamment pour ce qui est de la variété du vocabulaire et des tournures. Dans cette optique, il convient donc de tenir dûment compte de la discipline dont relève chaque travail de maturité.

La moyenne arithmétique de l'échelle de la qualité de la langue est de $M = 2,5$, avec un écart type de $SD = 0,91$. Sur l'ensemble des items, les appréciations des évaluateurs restent en moyenne positives. On n'observe pas de gros reproches en ce qui concerne la clarté de la langue et la continuité du texte. Il en va de même pour l'orthographe, la grammaire et la ponctuation. Au vu des moyennes, c'est l'expression langagière qui fait le score le plus faible: descriptions, analyses et réflexions sont souvent entrecoupées de tournures empruntées au langage familier.

Tableau VI.32: Valeurs moyennes et écarts types des critères de l'échelle «qualité de la langue»

Critères	1	2	3	4	5	6
Continuité du texte	2.2 (0.91)					
Clarté	2.6 (1.05)					
Choix des mots	2.5 (1.01)					
Expression	2.7 (1.10)					
Structure syntaxique	2.6 (0.96)					
Correction de la langue	2.3 (1.01)					

L'analyse de variance révèle à la comparaison multiple que la qualité de la langue des travaux de maturité est jugée significativement ($\alpha < 0,05$) meilleure en sciences de la nature ($M = 2,2$) qu'en sciences sociales ($M = 2,6$). En sciences humaines, on n'observe pas d'écarts significatifs d'un côté comme de l'autre. Il n'est pas impossible que la qualité de langue attribuée aux travaux de sciences de la nature et de mathématiques s'explique par les exigences différentes des experts à l'égard des travaux à caractère de propédeutique scientifique.

Le test t sur la distinction par le sexe ne fait pas apparaître d'écarts significatifs, ni non plus la comparaison entre travaux individuels et travaux d'équipe. Pour ce qui est de la comparaison régionale, l'analyse de variance montre une fois encore que les élèves romands ($M = 2,1$) tirent sensiblement mieux ($\alpha < 0,001$) leur épingle du jeu que leurs camarades alémaniques ($M = 2,8$). On retrouve au Tessin une moyenne très proche de celle des élèves romands ($M = 2,2$).

5.6 Qualité de la forme

La présentation écrite des résultats d'un travail scientifique fait l'objet de conventions qui changent d'une discipline à l'autre. On s'en rend compte, par exemple, en comparant les bibliographies publiées dans des revues de sciences humaines ou de sciences de la nature et de l'ingénieur. Mais à y regarder de plus près, on s'aperçoit qu'il existe aussi des variations en matière de présentation des citations ou des sources au fil du texte. Au sein d'une discipline, il importe aussi que la présentation soit uniforme, même si les consignes de présentation des citations et des références bibliographiques peuvent ne pas être les mêmes d'une publication à l'autre, comme c'est le cas en sciences sociales. L'évaluation de la qualité des travaux de maturité ne consiste pas à arbitrer entre les avantages et les inconvénients des conventions admises dans chaque domaine, mais à juger seulement de l'uniformité du respect des contraintes formelles.

La moyenne de l'échelle *qualité de la forme* est de $M = 2,5$, avec un écart type de $SD = 0,79$. La valeur moyenne des appréciations est donc positive, ce qui veut dire que les travaux se conforment en majorité aux exigences de présentation d'un texte à caractère de propédeutique scientifique. En ce qui concerne les divers critères, la page de titre et l'articulation formelle du texte sont en général jugées bonnes ou très bonnes par les évaluateurs; la plupart du temps, il suffit d'un rapide coup d'œil pour recueillir les principales informations sur le sujet, sur l'auteur ou les auteurs du travail, le contexte et le moment de sa préparation. Ces bons résultats s'expliquent notamment par la clarté des consignes émanant d'un très grand nombre d'écoles. Il en va de même pour la subdivision formelle du texte (introduction, corps et conclusion).

Les résultats sont tout juste positifs pour la présentation des citations et des sources et le traitement des tableaux et figures. Comme on l'a vu, les élèves ont surtout du mal à donner convenablement les références des sources trouvées sur l'Internet. D'une part des passages ne sont signalés que d'une façon globale, sans repérage précis de la citation; et d'autre part, les références ne sont fréquemment précisées que pour les reprises textuelles de passages; la paraphrase d'emprunts à des sources primaires ou secondaires ne fait souvent pas l'objet de références exactes. On observe aussi que l'appréciation portée sur les graphiques n'est pas aussi bonne chez les évaluateurs que chez les élèves eux-mêmes, qui manipulent moins bien qu'ils ne le croient l'outil informatique, par exemple pour ce qui est de leurs choix typographiques, des coupures de mots et de la numérotation des pages (voir section 4.5).

Tableau VI.33: Valeurs moyennes et écarts types de l'échelle «qualité de la forme»

Critères	1	2	3	4	5	6
Page de titre	1.8 (0.98)					
Présentation générale	2.5 (0.94)					
Visuels (tableaux et figures)	2.8 (1.12)					
Articulations formelles	2.0 (0.95)					
Présentation des citations	2.9 (1.41)					
Indication des sources	3.0 (1.56)					
Fonctionnalité des tableaux et fig.	2.8 (1.21)					

Pour la comparaison multiple après regroupement des travaux par disciplines, l'analyse de variance ne fait pas ressortir d'écarts significatifs entre les sciences humaines, les sciences sociales et les sciences de la nature. Le test t ne révèle pas de variations notables de résultats entre les travaux individuels ou en équipe pour ce qui est de la qualité formelle. Aucun écart significatif non plus entre les sexes. En revanche, la comparaison entre régions linguistiques montre que les élèves romands ($M = 2,3$) obtiennent des résultats significativement

meilleurs ($\alpha < 0,001$, bilatéral) que leurs camarades alémaniques ($M = 2,7$)⁶⁷ sur ce critère de la qualité de la forme, la moyenne des élèves italophones est de $M = 2,8$.

5.7 Appréciation globale

L'appréciation globale reflète l'impression générale que retire l'expert d'un travail de maturité. Dans la formation des évaluateurs, on a particulièrement insisté sur le fait qu'il ne s'agissait pas d'une note scolaire; mais l'échelle étant graduée de 1 à 6, la tentation restait forte de l'utiliser ainsi, même si sa polarité avait été inversée (le 1 représentant l'impression la meilleure et le 6 l'avis le plus défavorable, contrairement à l'usage scolaire suisse).

La moyenne arithmétique est de 3,1 pour l'appréciation globale, avec un écart type de 1,16. Le résultat moyen est ainsi juste en dessous de la moyenne virtuelle de l'échelle, c'est-à-dire légèrement positif. On peut donc dire que dans l'ensemble, les évaluateurs retirent des travaux une impression générale satisfaisante.

La corrélation est bonne entre l'appréciation globale et les trois échelles de qualité du fond, de la langue et de la forme. L'analyse de régression linéaire montre que 79% de la variance de l'appréciation globale s'expliquent par l'appréciation de la qualité du fond, de la langue et de la forme. C'est le fond qui a l'impact maximal, la forme qui compte le moins. Le coefficient bêta est de 0,58 pour la qualité du fond, de 0,23 pour celle de la langue, et de 0,1 pour celle de la forme.

Tableau VI.34: Corrélation entre l'appréciation globale et les échelles de qualité du fond, de la langue et de la forme

	Qualité du fond	Qualité de la langue	Qualité de la forme
Appréciation globale	0,85**	0,76**	0,62**

** La corrélation est significative au niveau 0,01 (bilatéral).

La comparaison multiple entre sciences humaines, sociales et de la nature ne fait pas apparaître de différence significative pour ce qui est de l'appréciation globale; le résultat est également négatif avec le test t sur le critère du sexe. Les travaux d'équipe ($M = 2,7$) jouissent d'une appréciation globale significativement meilleure ($\alpha < 0,05$, bilatéral) que les travaux réalisés en solitaire ($M = 3,1$). La comparaison régionale révèle que les élèves romands ($M = 2,6$) s'en tirent sensiblement mieux ($\alpha < 0,001$, bilatéral) que leurs camarades alémaniques ($M = 3,4$); au Tessin, cette valeur est de $M = 2,9$. Une réserve est toutefois de mise: faute d'évaluateurs multilingues, on ne saurait dire si ni dans quelle mesure c'est le reflet de cultures d'évaluation différentes qui apparaît ici.

5.8 Aspects individuels et institutionnels de la qualité

Nous avons déjà indiqué brièvement que l'appréciation des évaluateurs s'écarterait de celle des élèves eux-mêmes en ce qui concerne l'utilisation du traitement de texte et la présentation graphique du rapport du travail de maturité. Nous allons maintenant examiner systématiquement l'évaluation faite de la qualité des travaux de maturité par leurs auteurs.

Choix du sujet

Nous avons fait l'hypothèse que la qualité de fond devait aller en s'améliorant avec la liberté de choix du sujet. L'analyse de variance avec correction de Bonferroni pour comparaison multiple montre au contraire que la qualité du fond des travaux tombe significativement ($\alpha < 0,05$) chez les élèves qui choisissent librement leur sujet

⁶⁷ La comparaison régionale ne tient pas compte de l'échantillon tessinois en raison du nombre réduit de cas pour un nombre relativement important de spécialistes assurant l'évaluation.

(M=2,8) par rapport à ceux qui doivent le sélectionner dans une liste (M=2,2) ou à l'intérieur d'un thème général (M=2,4). Il apparaît par ailleurs que l'aide spécifiquement apportée par l'enseignant au choix du sujet a un effet positif sur la qualité du contenu du travail. Cela pourrait s'expliquer par l'encadrement: il est plausible que l'enseignant parvienne mieux à réagir à des questions et à des problèmes spécifiques lorsque le choix du sujet a été structuré et limité d'avance, même si les personnes interrogées n'en avaient pas conscience. La perception de l'accompagnement dont ils ont bénéficié est la même chez les élèves, quel que soit le degré de liberté qu'ils ont eu dans le choix de leur sujet. La meilleure qualité de fond des travaux des élèves dirigés dans le choix de leur sujet donne à penser que l'encadrement a pu être plus efficace dans ce cas. Cette interprétation est renforcée par le fait que l'on n'observe pas de lien significatif entre la perception de l'accompagnement et la qualité des travaux ressortant de leur évaluation.

Les restrictions au choix du sujet peuvent aussi avoir des effets positifs sur la coopération entre les élèves. Les échanges seront plus faciles entre eux s'ils travaillent sur des questions proches que lorsque les sujets sont très éloignés, dans la classe ou l'année.

Il est relativement étonnant d'observer que la liberté de choix du sujet ne se traduit pas automatiquement par une meilleure motivation ni de meilleurs résultats, ce qui souligne l'importance d'un cadre institutionnel impératif et clairement formulé. Nos résultats nous apprennent que ces deux choses ont plus d'impact sur la qualité de fond des travaux que l'intérêt ou les connaissances préalables des élèves, qui n'ont pas de répercussions significatives sur elle.

Aucune variation significative de qualité n'a été observée entre les travaux relevant des sciences humaines, sociales et de la nature. On constate par ailleurs que les travaux interdisciplinaires ne présentent pas non plus de différence perceptible avec des travaux plus clairement rattachés à une matière.

Modes opératoires

Comme pour la perception individuelle de l'encadrement, la fréquence des contacts entre l'élève et la personne assurant son suivi ne semble pas avoir d'impact significatif sur les divers aspects de la qualité des travaux. On remarque toutefois certains points d'une grande importance pour l'élève dans la préparation du travail de maturité. Outre l'aide dispensée dans le choix du sujet, l'assistance fournie sur le maniement des médias électroniques, sur les méthodes de travail, sur le respect des règles de forme et sur la rédaction du rapport à eu des effets. L'aide à la rédaction se reflète non seulement sur le contenu, mais aussi sur la qualité de la langue, et contribue à celle de la forme du travail. Mais ces corrélations sont en général très ténues.

Le tableau VI. 35 présente les corrélations avec les modes opératoires. L'assistance a été analysée en divers items, mesurés sur une échelle à quatre degrés de 1 (très utile) à 4 (inutile). La polarité a été inversée, comme dans la grille d'évaluation.

Tableau VI.35: Corrélation entre l'appréciation portée sur l'assistance reçue de l'enseignant suivant le travail et les échelles de qualité

	Qualité fond	Qualité langue	Qualité forme
Choix du sujet	0,14*	non significatif	non significatif
Problématique	non significatif	non significatif	non significatif
Gestion du temps	non significatif	non significatif	non significatif
Outil informatique	0,14*	non significatif	non significatif
Contenu (fond)	non significatif	non significatif	non significatif
Méthodes de travail	0,11*	non significatif	non significatif
Structure, articulations	non significatif	non significatif	non significatif

	Qualité fond	Qualité langue	Qualité forme
Rédaction du travail	0,14*	0,16**	non significatif
Motivation	non significatif	non significatif	non significatif
Forme	0,15**	0,13*	0,15**

* La corrélation est significative au niveau 0,05 (bilatéral).

** La corrélation est significative au niveau 0,01 (bilatéral).

La qualité des travaux des élèves qui ont assisté à des séances de préparation est en général légèrement meilleure pour le fond et la langue, mais pas pour la forme. On trouve en particulier des corrélations significatives avec les séances d'information et l'apprentissage sur projet. Ce dernier est toutefois une forme de cours très spéciale, à laquelle ne recourent qu'un tout petit nombre d'élèves (32 dans notre échantillon). Nos résultats montrent en tout cas qu'il est utile d'organiser des séances formelles de préparation et d'accompagnement des travaux de maturité. Les élèves qui indiquent n'avoir assisté à aucune tout en les jugeant utiles atteignent dans leurs travaux une qualité de langue inférieure, sans que leurs résultats soient significativement moins bons en ce qui concerne la forme et le fond. Le groupe de référence est formé dans ce cas de tous les élèves interrogés qui n'ont pas suivi de séances de préparation et ne pensent pas avoir été handicapés pour autant.

Il n'y a aucun rapport entre le temps que consacrent les élèves à leur travail et la qualité de ce dernier. Cela vaut pour le temps que les élèves déclarent lui avoir consacré dans leurs questionnaires comme pour celui que prescrivent les écoles. On ne trouve pas non plus de corrélation entre la qualité des travaux et les connaissances préalables que les élèves estiment avoir possédées. Cependant, les élèves qui jugent avoir consacré un effort excessif à leur travail par rapport à d'autres tâches ont obtenu une évaluation un peu moins bonne à leur rapport de projet ($r=0,13^{**}$ pour la qualité de la langue et $r=0,11^{*}$ pour celle du fond).

L'existence d'un guide et la transparence des procédures aux yeux des élèves n'ont pas d'impact positif sur la qualité des travaux de maturité, qui semble dépendre bien davantage du suivi pédagogique direct. La connaissance des critères d'évaluation ne paraît pas non plus avoir d'effet positif; les travaux dont les auteurs déclarent n'en avoir pas eu connaissance à l'avance sembleraient même avoir été jugés plutôt meilleurs pour ce qui est de la qualité de la langue ($r=0,15^{**}$). On observe aussi une faible corrélation négative entre les appréciations portées par les élèves sur le bien-fondé des critères et la qualité de la langue et du fond de leurs travaux ($r=-0,17^{**}$ pour la langue, et $r=-0,11^{*}$ pour le fond). Les élèves qui trouvent explicitement les critères injustifiés ont aussi tendance à produire des travaux de moins bonne qualité. Ce résultat risque là encore être le produit d'une réaction *a posteriori*: les élèves n'ayant pas très bien réussi leur travail de maturité pourraient mettre en doute la valeur des critères d'appréciation.

Outre l'aide de l'enseignant, d'autres personnes ont aussi par leur appui une influence sur la qualité des travaux de maturité. Les élèves obtiennent en général une meilleure appréciation de qualité formelle de leur travail de maturité ($r=0,11^{*}$) lorsqu'ils se sont fait aider par des frères et sœurs, et une meilleure appréciation pour le contenu de leur travail ($r=0,1^{*}$) quand ils ont bénéficié du concours de camarades d'école. Ces corrélations sont indéniablement faibles, mais significatives. En revanche, l'aide des parents ne semble pas avoir de répercussions notables sur la qualité du travail. On remarquera toutefois dans ce contexte qu'il ne faut pas trop attendre de la déclaration des élèves à ce sujet, qui ne dit rien de l'influence latente des parents sur l'itinéraire de formation de l'élève.

Utilité du travail de maturité

De l'avis des élèves, leur travail de maturité a eu des effets positifs sur divers aspects de leur façon d'organiser leur travail et de leurs compétences cognitives (voir section 4). Les auteurs de travaux présentant une bonne qualité de langue reconnaissent que le projet a eu des effets bénéfiques sur leurs compétences cognitives ($r=0,12^{*}$). A part cette corrélation très faible, on ne détecte aucun effet significatif entre la qualité des travaux et

leur utilité perçue dans l'organisation du travail et les compétences cognitives. La note donnée aux résultats du projet présente en revanche une étroite corrélation avec l'utilité perçue du travail de maturité ($r=0,23^{**}$ pour l'effet sur les compétences cognitives, et $r=0,3^{**}$ pour l'effet sur la capacité de s'organiser dans son travail). Une fois encore, l'évaluation des élèves semble avoir été assez nettement influencée par l'appréciation portée par l'institution sur leur travail de maturité.

On observe une légère corrélation entre la qualité de la langue et du fond des travaux et l'intérêt accru mentionné par leurs auteurs ($r=0,11^*$ dans l'un et l'autre cas). Le regain d'intérêt, en revanche, présente une corrélation beaucoup plus forte avec la note ($r=0,3^{**}$). La corrélation est nette aussi entre le plaisir éprouvé à préparer le travail et la note ($r=0,34^{**}$), mais il n'y en a pas de significative avec le résultat de l'évaluation. Les données dont nous disposons ne nous permettent pas de déterminer si le succès a teinté *a posteriori* la perception de l'intérêt et du plaisir éprouvés en cours de travail, ou si ce sont au contraire l'intérêt et la motivation qui se reflètent dans la note.

On remarquera que les élèves ayant obtenu dans l'évaluation une appréciation supérieure à la moyenne ont une plus haute opinion de l'utilité de leur travail de maturité pour leurs études supérieures, leur formation professionnelle et leur développement personnel. L'échelle utilisée comportait à chaque fois quatre degrés, de 1 (très réduite) à 4 (très grande); le tableau VI.36 présente les corrélations correspondantes.

Tableau VI.36: Corrélation entre l'utilité perçue et les trois échelles d'appréciation de la qualité du fond, de la langue et de la forme

Utilité	Qualité fond	Qualité langue	Qualité forme
Etudes supérieures	0,12*	non significatif	non significatif
Formation professionnelle	0,12*	0,10*	non significatif
Développement personnel	0,11*	non significatif	0,12*

* La corrélation est significative au niveau 0,05 (bilatéral).

5.9 Bilan partiel

L'analyse confirme, sur la base de l'évaluation, la validité de trois échelles de mesure de la qualité du fond, de la forme et de la langue des travaux – trois grands aspects de la propédeutique scientifique. Les valeurs moyennes des échelles étant toutes sur le versant positif, la qualité peut être globalement qualifiée de satisfaisante. Si l'évaluation donne des résultats nettement en dessous des notes scolaires attribuées aux travaux, c'est surtout parce que le jugement pédagogique inclut des aspects formatifs que ne prend pas en considération l'évaluation par simple addition effectuée dans l'optique de la propédeutique scientifique.

Les échelles d'appréciation de la qualité du fond, de la forme et de la langue saturent chacune sur un facteur, sans être complètement indépendantes les unes des autres dimensionnellement. Ce qui apparaît dans le fait que la corrélation est positive entre elles, mais aussi entre chacune et l'appréciation globale reflétant l'impression générale que retire l'évaluateur du travail (et qui se rapproche d'une note scolaire, puisqu'elle est exprimée sur une échelle à six degrés).

Au niveau des cadres institutionnels, de nombreux indices conduisent à penser que les contraintes resserrant le choix du sujet favorisent mieux la qualité des travaux que le libre choix. Ce qui veut dire que notre hypothèse de causalité entre le libre choix du sujet et la motivation intrinsèque, et ainsi la qualité du travail, ne s'est pas vérifiée. Les consignes visant à structurer les phases du projet apparaissent donc comme un bon atout, et leur caractère d'orientation pourrait être renforcé par la constatation qu'elles consolident les liens entre les élèves et permettent aux enseignants de faire un emploi plus concentré de leurs ressources.

Les élèves ne prennent qu'assez accessoirement connaissance des guides et autres documents officiels. La transparence se fait surtout pour eux par les contacts personnels avec les enseignants qui les suivent. Le sérieux avec lequel ces derniers assument leur rôle est indéniablement un important facteur de succès, tandis que ni la fréquence des contacts entre eux et les élèves, ni la longueur de la période de préparation ne semblent jouer de grand rôle dans le succès de l'entreprise.

Ce sont surtout les auteurs de travaux d'une qualité supérieure à la moyenne qui portent une appréciation positive sur les apports de l'exercice en termes de qualifications universitaires et professionnelles et de développement personnel. Mais la corrélation ne permet pas à non plus de conclure au lien de causalité: l'utilité attendue peut avoir une influence sur la qualité des résultats; et inversement, un bon résultat peut influencer *a posteriori* le jugement que l'on porte sur les apports présumés de l'exercice.

6 Conclusions

Ce projet a donné l'occasion d'évaluer plus de quatre cents travaux de maturité de Suisse alémanique, de Suisse romande et du Tessin. Le résultat général est positif: la qualité de propédeutique scientifique de ces travaux est en majeure partie satisfaisante. On remarque en même temps que notre évaluation a été en général plus critique que celle des enseignants. Il ne faut toutefois pas oublier que le centrage de notre étude diffère sur bien des aspects de celui d'une évaluation pédagogique. Les enseignants ne se prononcent d'habitude pas sur le seul produit écrit, mais aussi sur la présentation orale. D'autre part, notre analyse de qualité a procédé par pure addition, car nous n'avons pas pu accompagner le processus; les enseignants, en revanche, ont eu l'occasion d'observer par eux-mêmes les progrès des élèves en cours de projet. Il faut par ailleurs souligner que l'évaluation à laquelle nous avons procédé se limitait en grande partie à la perspective académique de l'aptitude aux études supérieures, alors que les enseignants sont conduits par leur mission éducative à intégrer un grand nombre d'autres facteurs dans leur appréciation.

L'analyse a confirmé l'hypothèse de corrélation entre la qualité de fond des travaux et leur qualité de forme et de langue: les travaux présentant une bonne qualité de fond obtiennent en général aussi de bons scores sur les autres critères; ceux dont la qualité de fond est médiocre souffrent en outre d'insuffisances de forme et de langue. Les résultats sont meilleurs lorsque les élèves doivent choisir leur sujet dans un thème général imposé, un peu moins bons lorsqu'ils le définissent tout à fait librement – ce qui pourrait s'expliquer moins par la liberté de choix que par la plus grande complexité du suivi par laquelle elle se traduit.

L'analyse révèle le rôle fondamental des contacts interpersonnels entre élèves et enseignants dans le succès de l'entreprise. Ce qui compte, ce sont moins la nature et la fréquence des contacts que l'intensité perçue du soutien dans les phases critiques. Cela commence avec la formulation du sujet et va jusqu'à la structuration formelle du rapport, en passant par la clarification des questions méthodologiques. La longueur du temps imparti à la préparation du projet a aussi peu d'impact sur la qualité du travail de maturité que la fréquence des contacts avec l'enseignant.

L'efficacité de l'encadrement grandit lorsque ce n'est pas l'école qui répartit arbitrairement les élèves entre les enseignants, mais les élèves eux-mêmes qui choisissent l'enseignant qui les suivra, notamment sur le critère de la sympathie personnelle. Outre les enseignants, on note aussi le rôle de l'entourage social des élèves, même s'il est étonnant de constater que ses effets ne sont faibles mais significatifs que pour les frères et sœurs et les camarades. L'influence des parents évoquée par les élèves n'a pas de retombées perceptibles sur la qualité des travaux; encore faut-il souligner que l'on ne dispose que de peu de données sur l'entourage des élèves, et qu'il n'a été tenu aucun compte de l'influence latente des parents sur l'itinéraire de formation de leurs enfants.

Les résultats généralement meilleurs pour ce qui est de la langue et du contenu observés dans les travaux réalisés en équipe sont probablement imputables en bonne part à la communication au sein du groupe. On retrouve ici l'influence des pairs sur la qualité du travail. Ce phénomène disparaît pour ce qui est de la qualité de la forme, qui dépend surtout de l'efficacité institutionnelle des consignes données dans le cadre du suivi pédagogique.

L'évaluation montre que, contre toute attente, la qualité des travaux de maturité n'est pas très influencée par l'intérêt personnel ou les connaissances spécifiques préalables que les élèves pensent après coup avoir possédées. On observe une corrélation positive faiblement significative entre l'accroissement perçu par les élèves de leurs compétences cognitives et la qualité de la langue et du fond des travaux; il n'est pas étonnant par ailleurs que l'appréciation portée *a posteriori* à ce sujet dépende plus étroitement de la note obtenue.

L'impression de transparence des procédures, des objectifs et de la notation n'a pas d'effets significatifs sur la qualité des travaux. Les élèves ne prennent qu'assez accessoirement connaissance des manuels et guides préparés à leur intention par les écoles, sans que cela veuille dire que ces documents sont inutiles: c'est plutôt qu'ils jouent un rôle indirect, par le biais de l'encadrement du projet. De même, les séances et cours de préparation n'ont que des répercussions modestes sur les divers volets de la qualité. On note les bénéfices

significatifs des séances d'information spéciales et de l'expérience acquise dans le cadre de l'apprentissage par projet; plus leur utilité est appréciée, meilleure est la qualité de fond du travail. Les élèves qui regrettent l'absence de séances préparatoires tirent un peu moins bien leur épingle du jeu en ce qui concerne la qualité de leurs travaux.

Les élèves ont tendance à distinguer une utilité intrinsèque dans leur travail de maturité; ils sont relativement rares à le juger inutile par la suite. Plus de la moitié en attendent des bénéfices pour leurs études supérieures, nettement moins de la moitié pour une formation professionnelle. La corrélation est positive entre la qualité de fond et l'avis des élèves sur l'utilité de leur travail de maturité en termes de qualifications académiques ou professionnelles, ainsi que de développement personnel.

Il n'y a pas de grosses différences de qualité entre les travaux de maturité relevant des sciences humaines, sociales ou de la nature. Le critère du sexe ne fait pas non plus apparaître d'écart sensible. La comparaison régionale, en revanche, en révèle: si le résultat est positif dans l'ensemble du pays, on observe des variations importantes d'une région linguistique à l'autre. Les Romands font un meilleur score que les Alémaniques; les Tessinois, tout en s'en sortant mieux que leurs camarades alémaniques, sauf sur le critère de la qualité formelle, font un peu moins bien dans l'ensemble que leurs camarades romands dans l'évaluation. Ces écarts n'ont pas pu être statistiquement expliqués dans notre projet. Nous avions l'intention de recourir à un grand nombre d'évaluateurs bilingues qui auraient pu examiner des travaux de maturité dans diverses disciplines scientifiques à travers les frontières régionales; mais nous n'avons pas pu le faire, pour des raisons d'organisation. Ce qui veut dire que nous n'avons pas disposé d'étalon de comparaison entre la Suisse romande et la Suisse alémanique (la comparaison n'est pas vraiment possible avec le Tessin, compte tenu de l'exiguïté de l'échantillon). On ne saurait donc exclure totalement que malgré l'uniformité des instruments utilisés et de la formation dispensée aux évaluateurs, on retrouve ici l'effet de cultures d'évaluation différentes, se traduisant par des écarts d'exigences d'une région linguistique à l'autre. Mais un certain nombre d'indices conduisent à penser que le meilleur score des élèves romands par rapport à leurs camarades alémaniques ne remonte pas à des particularités régionales, mais à des facteurs institutionnels. Et il faut enfin signaler que l'organisation du travail de maturité au Tessin est sensiblement plus proche de ce qui se fait en Suisse romande qu'en Suisse alémanique.

Aux yeux des établissements d'enseignement supérieur, les bacheliers devraient avoir appris par leur travail de maturité à répondre aux exigences élémentaires des études qui les attendent: par la structuration du texte, la présentation des citations et de la bibliographie, le travail de maturité témoigne que le gymnase a familiarisé les élèves avec les règles courantes du travail universitaire. Il en va de même pour la présentation de théories ou de faits empiriques élémentaires, ainsi que pour les premiers efforts de collecte d'informations par la recherche et l'expérimentation – même si le travail descriptif et reproductif l'emporte nettement sur l'analyse dans les projets scolaires.

Il convient d'ailleurs d'émettre à ce propos une réserve importante: le travail de maturité est une étape finale unique en son genre dans l'acquisition par l'élève de la maturité intellectuelle universitaire. Le caractère original du projet, la variété des dispositifs institutionnels, l'intensité de l'encadrement pédagogique et l'appui de l'entourage social sont autant de facteurs contextuels dont la conjonction contribue à la qualité du produit final. Si cette conjonction ne se reproduit pas à l'université, on peut s'attendre à ce que les premiers travaux rendus au cours des semestres initiaux restent souvent en dessous du niveau du travail de maturité. Il ne nous a pas été possible, dans le cadre de notre projet, de vérifier le caractère durable des acquis obtenus par les élèves à l'occasion de leur travail de maturité. Il n'est pas possible d'en tirer de prévisions sur la mesure dans laquelle l'expérience ainsi accumulée est transférable à une formation universitaire ou professionnelle après le gymnase.

Même si ce caractère durable n'est pas prouvé, de nombreux indices portent à croire que le travail de maturité est un outil pédagogique judicieux et fécond. Les résultats de notre évaluation comme les réponses des élèves incitent à maintenir au gymnase la méthodologie du projet, et à l'étoffer avec modération.

Références bibliographiques du chapitre VI

Sources du chapitre IV

- Argovie: Verordnung über die Promotion und die Maturität an den Mittelschulen (Maturitätsverordnung) vom 23. Juni 1999.
- Berne: Ordonnance de Direction du 3 juillet 1997 concernant l'accès à la formation gymnasiale et l'enseignement dans les écoles de maturité (ODEMa).
- Berne: Ordonnance du 27 novembre 1996 sur les écoles de maturité (OEMa).
- Bâle-Campagne: Verordnung über die Maturitätsprüfungen vom 5. Juli 2005.
- Bâle-Campagne: Weisungen der Rektorenkonferenz zu den Maturaarbeiten vom 9. Mai 2000.
- Bâle-Ville: Verordnung betreffend die Maturitätsprüfungen im Kanton Basel-Stadt vom 28. März 2000.
- Bâle-Ville: Règlement für die Maturaarbeit an den Gymnasien im Kanton Basel-Stadt (Règlement für die Maturaarbeit) vom 2. Februar 2001.
- Fribourg: Règlement du 15 avril 1998 sur les études gymnasiale (REG).
- Fribourg: Règlement du 17 septembre 2001 concernant les examens de baccalauréat (REB).
- Fribourg: Lignes directrices de la Direction de l'instruction publique, de la culture et du sport du 21 janvier 2004 concernant la réalisation des travaux de maturité.
- Genève: Règlement du 14 octobre 1998 relatif à la formation gymnasiale au collège de Genève.
- Grisons: Verordnung über das Gymnasium im Kanton Graubünden (GymVO) vom 6. Juli 1999.
- Jura: Règlement du 16 janvier 1995 sur la reconnaissance des certificats de maturité gymnasiale (RRM).
- Lucerne: Règlement für die Maturitätsprüfungen im Kanton Luzern vom 27. Mai 1999.
- Lucerne: Die Maturaarbeit an den Luzerner Maturitätsschulen; Weisungen der Maturitätskommission vom 11. Januar 2000, revidiert am 4. Juli 2000.
- Neuchâtel: Règlement du 13 mai 1997 des études des lycées cantonaux (admission, promotion et examens).
- Obwald: Règlement über die Maturitätsprüfungen vom 22. April 1997.
- Obwald: Rahmenbestimmungen für die Maturaarbeit vom 22. April 2002.
- Saint-Gall: Maturitätsprüfungsreglement des Gymnasiums vom 24. Juni 1998.
- Thurgovie: Verordnung des Regierungsrates über die Maturitätsabteilung an den thurgauischen Kantonsschulen vom 6. Juli 1999.
- Tessin: Regolamento degli studi liceali (del 24 giugno 1997).
- Valais: Règlement du 10 avril 2002 concernant les études gymnasiales et les examens de maturité.
- Valais: Directives du 10 décembre 2004 relatives au travail de maturité gymnasiale.
- Vaud: Règlement des gymnases (RGY) du 7 mai 1997.
- Vaud: Loi du 17 septembre 1985 sur l'enseignement secondaire supérieur (LESS).

Zoug: Reglement über die Maturitätsprüfungen an den kantonalen und privaten Gymnasien vom 17. März und 21. August 1997.

Zurich: Reglement für die Maturitätsprüfungen an den Gymnasien des Kantons Zürich vom 10. März 1998.

Bibliographie du chapitre VI

Arbeitsstelle für Hochschuldidaktik (2006): *Leistungsnachweise in modularisierten Studiengängen*. Dossier. Zurich: Universität Zürich, Arbeitsstelle für Hochschuldidaktik.

Bieri Buschor, Christine & Forrer, Esther (2005): Cool, kompetent und kein bisschen weise? Überfachliche Kompetenzen junger Erwachsener am Übergang zwischen Schule und Beruf. Zurich, Coire: Verlag Rüegger (=Wissenschaftliche Reihe, Bd. 18).

Binder, Hans-Martin (2003): Maturaarbeit - Engagement und Belastung. Ergebnisse einer Maturandenbefragung in Luzern. Dans: *Neue Zürcher Zeitung*, Nr. 109, 13.5.2003, p. 69.

Binder, Hans-Martin & Feller-Länzlinger, Ruth (2004): *Maturitätsreform im Kanton Aargau. Schlussbericht der externen Evaluation*. Lucerne: Interface, Institut für Politikstudien.

Binder, Hans-Martin & Feller-Länzlinger, Ruth (2003): Maturaarbeit und Maturaprüfung nach neuem Maturitätsanerkennungsreglement MAR. Auswertung der schriftlichen Befragung der AbsolventInnen der Matura 2002 im Rahmen der externen Evaluation der Gymnasialreform im Kanton Luzern. Lucerne: Interface, Institut für Politikstudien.

Bonati, Peter (2007): Was Gymnasium und Standards voneinander lernen können. Dans: Peter Labudde (éd.), *Bildungsstandards am Gymnasium. Korsett oder Katalysator?* Berne: hep, p. 47-56.

Bonati, Peter (2006): Gymnasium wohin? - Bojen für die Weiterfahrt. Dans: *Gymnasium Helveticum*, 3, p. 5-15.

Bonati, Peter & Hadorn, Rudolf (2007): Matura- und andere selbstständige Arbeiten betreuen. Ein Handbuch für Lehrpersonen und Dozierende. Berne: hep.

Carolus, Thomas (2006): Hinweise zur Anfertigung einer Studien- oder Diplomarbeit. Siegen: Universität Siegen, Fachbereich Maschinenbau, Institut für Fluid- und Thermodynamik. http://www.uni-siegen.de/fb11/iftsm/aktuelles/anleitungwissarbeiten_3.pdf (Etat: 15.9.2008)

CDIP [Conférence suisse des directeurs cantonaux de l'instruction publique] (2007): *Formation gymnasiale: plus de poids accordé aux sciences expérimentales et au travail de maturité*. Communiqué de presse du 27.6.2007. Berne: CDIP.

CDIP [Conférence suisse des directeurs cantonaux de l'instruction publique] (éd.) (2000): Le secondaire II à venir. Rapport final du groupe de projet secondaire II (CDIP/OFFT). Berne: CDIP.

CDIP [Conférence suisse des directeurs cantonaux de l'instruction publique] (1995): *Ordonnance du Conseil fédéral/Règlement de la CDIP des 16 janvier/15 février 1995 sur la reconnaissance des certificats de maturité gymnasiale (RRM)*.

CDIP [Conférence suisse des directeurs cantonaux de l'instruction publique] (1994): Plan d'études cadre du 9 juin 1994 pour les écoles de maturité. Recommandation à l'intention des cantons conformément à l'art. 3 du Concordat scolaire du 29 octobre 1970. Avec des propositions pour sa mise en œuvre. Berne: CDIP.

- CDIP [Conférence suisse des directeurs cantonaux de l'instruction publique] (éd.) (1972):
Enseignement secondaire de demain. *Annuaire de la Conférence suisse des directeurs cantonaux de l'instruction publique*, tome 58.
- Criblez, Lucien (2005): Das Gymnasium im Stress. Dans: *amv-aktuell*, Sonderheft 2005/1, p. 4-12.
- Darbellay, Vincent (2002): Travail de maturité: Une esquisse de bilan. Dans: Kollegiumsbote - Le Message du Collège, H. 1, p. 12-14.
- Davaud, Clairette & Hexel, Dagmar (2003): La nouvelle maturité gymnasiale. Orientations et intérêts des élèves de 4^e. Genève: Service de la Recherche en Education.
<http://www.geneve.ch/sred/publications/docsred/2003/NouvMatuGym.pdf> (Etat: 19.12.2007).
- Deutscher Bildungsrat (1970): Empfehlungen der Bildungskommission: Strukturplan für das deutsche Bildungswesen. Stuttgart: Klett.
- Dreyer, Hans Peter (2005): Gymnasium und Universität: entfremdete Geschwister. Dans: *amv-aktuell*, Sonderheft 2005/1, p. 22-27.
- Dubs, Rolf (1996): Innovation und Qualität am Gymnasium. Dans: Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP) (éd.), *De l'«enseignement secondaire de demain» à la réforme de la maturité 1995*. Berne: CDIP, p. 117-133.
- Eberle, Franz (2007): Zum Fundament von Bildungsstandards an Schweizer Gymnasien. Dans: Peter Labudde (éd.), *Bildungsstandards am Gymnasium. Korsett oder Katalysator?* Berne: hep, p. 37-45.
- Eberle, Franz (2005): EVAMAR Phase II – Eine Übersicht über das Evaluationsprojekt. Dans: *Gymnasium Helveticum*, 6, p. 5-11.
- Erdfelder, Edgar/Jochen Musch & Lutz Cüpper (2002): *Richtlinien zur Gestaltung von schriftlichen Referaten, Hausarbeiten und Praktikumsberichten*. Mannheim: Universität Mannheim, Lehrstuhl Psychologie III. <http://www.psychologie.uni-mannheim.de/psycho3/Sonstiges/Richtlinien.pdf> (Etat: 15.9.2008).
- Fontollet, Pierre-Gérard (1996): Une maturité de demain – Objectifs et réalités d'une réforme. Dans: Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP) (éd.), *De l'«enseignement secondaire de demain» à la réforme de la maturité 1995*. Berne: CDIP, p. 47-60.
- Frey, Karl (1993): *Die Projektmethode*. Weinheim: Beltz.
- Furrer, Hans (2000): Ressourcen – Kompetenzen – Performanz. Kompetenzmanagement für Fachleute der Erwachsenenbildung. Lucerne, Zurich: Akademie für Erwachsenenbildung AEB.
- Grin, François; Alliata, Roberta; Dozio, Edo; Labudde, Peter; Maag Merki, Katharina & Pagnossin, Elisabetta (2004): Evaluation de la réforme de la maturité 1995 (EVAMAR). Document supplémentaire de la partie 3. Module 2. Objectifs pédagogiques transversaux. Genève: Service de la recherche en éducation (SRED).
- Habel, Werner (1990): Wissenschaftspropädeutik. Untersuchungen zur gymnasialen Bildungstheorie des 19. und 20. Jahrhunderts. Cologne, Vienne: Böhlau.
- Hadorn, Rudolf (2004): *Unterlagen zum Atelier 6: Maturaarbeit*. [Notes de présentation non publiées, distribuées dans le cadre du séminaire pour cadres «Matura light» du 25.8.2004 à Berne]
- Heldmann, Werner (1998): Studieren heute. Erwartungen der einzelnen Studienfächer an ihre Studienanfänger. (2., völlig Neubearb. Aufl.) Bad Honnef: Bock.
- Hentig, Hartmut von (1996): *Bildung. Ein Essay*. Munich, Vienne: Hanser.

- Hirschi, Hans (2002): Lehrpläne vermutlich noch zu ehrgeizig. Dans: *Mitteilungsblatt Kanton Luzern*, 5, p. 29-30.
- Huber, Ludwig (1998): Allgemeine Studierfähigkeit, basale Fähigkeiten, Grundbildung. Zur aktuellen Diskussion um die gymnasiale Oberstufe. Dans: Rudolf Messner, Erhard Wicke & Dorit Bosse (éd.), *Die Zukunft der gymnasialen Oberstufe. Beiträge zu ihrer Weiterentwicklung*. Weinheim, Bâle: Beltz, p. 150-181.
- Huber, Ludwig (1997): Fähigkeit zum Studieren - Bildung durch Wissenschaft? Zum Problem der Passung zwischen Gymnasialer Oberstufe und Hochschule. Dans: Eckart Liebau, Wolfgang Mack & Christoph Scheilke (éd.), *Das Gymnasium. Alltag, Reform, Geschichte, Theorie*. Weinheim, Munich: Juventa, p. 333-351.
- Hupka, Sandra & Dellenbach, Myriam (2004): *Konsequenzen der Maturitätsreform 1995 für das Bestehen der Matura. Ergebnisse der Notenerhebung im EVAMAR-Projekt*. Berne: Erziehungsdirektion des Kantons Bern, Bildungsplanung und Evaluation.
- Ingenkamp, Karlheinz (éd.) (1989). *Die Fragwürdigkeit der Zensurenggebung: Texte und Untersuchungsberichte*. Weinheim, Bâle: Beltz.
- Inglis, Oswald (2007): Macht die Maturaarbeit nicht kaputt! Dans: *Gymnasium Helveticum*, 1, p. 27-28.
- Jacob, Walter (2004): Maturaarbeiten: Tut doch endlich etwas! Dans: *Gymnasium Helveticum*, 5, p. 5-8.
- Köller, Olaf (2007): Bildungsmonitoring und Vergleichsarbeiten in Deutschland. Dans: Peter Labudde (éd.), *Bildungsstandards am Gymnasium. Korsett oder Katalysator?* Berne: hep, p. 65-73.
- Köller, Olaf; Watermann, Rainer; Trautwein, Ulrich & Lüdtke, Oliver (2004): Wege zur Hochschulreife in Baden-Württemberg - Erweiterung von Bildungswegen und Studiereignung: Die grundlegenden Fragestellungen in TOSCA. Dans: Olaf Köller, Rainer Watermann, Ulrich Trautwein & Oliver Lüdtke (éd.): *Wege zur Hochschulreife in Baden-Württemberg. TOSCA - Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien*. Opladen: Leske + Budrich, p. 113-119.
- Klieme, Eckhard et al. (2003): *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise* (2. Auflage). Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- KMK [Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland] (2006): Vereinbarung zur Gestaltung der gymnasialen Oberstufe in der Sekundarstufe II. Beschluss der Kultusministerkonferenz vom 7.7.1972 i.d.F. vom 2.6.2006.
<http://www.kmk.org/doc/publ/Vereinb-z-Gestalt-d-gymOb-i-d-SekII.pdf> (Etat: 27.2.2008).
- KMK [Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland] (1978): Empfehlungen zur Arbeit in der gymnasialen Oberstufe. Dans: *Bildung und Erziehung*, 31, p. 561-574.
- Konegen-Grenier, Christiane (2002): *Studierfähigkeit und Hochschulzugang*. Cologne: Deutscher Instituts-Verlag.
- Kruse, Otto (1999): *Keine Angst vor dem leeren Blatt. Ohne Schreibblockaden durchs Studium*. (7. Auflage.) Francfort et New York: Campus.
- Kunz, Christin M. (2003): Gesellenstück gymnasialer Bildung. Maturaarbeit auf der zweiten Stufe gymnasialer Reform. Dans: *Neue Zürcher Zeitung*, 13.5.2003, p. 69.
- Linke, Angelika (2006): Die Matura in den Zeiten von Bologna. Dans: *Gymnasium Helveticum*, 4, p. 6-10.
- Loos, Barbara & Popp, Susanne (1998): Praxis der gymnasialen Oberstufe. Varianten fachübergreifenden und fächerverbindenden Lernens und Arbeitens. Dans: Rudolf Messner,

- Erhard Wicke & Dorit Bosse (éd.), *Die Zukunft der gymnasialen Oberstufe. Beiträge zu ihrer Weiterentwicklung*. Weinheim, Bâle: Beltz, p. 126-149.
- Lorenzen, Klaus F. (2002): *Wissenschaftliche Anforderungen an Diplomarbeiten und Kriterien ihrer Beurteilung*. (3., völlig neu bearb. Ausg. 10.2.2002, zuerst ersch. am 8.1.1997). <http://bui.haw-hamburg.de/pers/klaus.lorenzen/ASP/wisskrit.pdf> (Etat: 4.12.2006).
- Maag Merki, Katharina (2006): Teil III: Bilanz. Dans: Katharina Maag Merki (éd.): *Lernort Gymnasium. Individuelle Entwicklungsverläufe und Schulerfahrungen*. Berne: Haupt, p. 175-194.
- Maag Merki, Katharina & Leutwyler, Bruno (2006): Die Fähigkeit zum selbst regulierten Lernen. Dans: Katharina Maag Merki (éd.), *Lernort Gymnasium. Individuelle Entwicklungsverläufe und Schulerfahrungen*. Berne: Haupt, p. 79-137.
- Meylan, Jean-Pierre (1996): Die Erneuerung des Gymnasiums und die Anerkennung der Maturitäten. Stationen der Debatte 1968-1995. Dans: Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP) (éd.), *De l'«enseignement secondaire de demain» à la réforme de la maturité 1995*. Berne: CDIP, p. 7-45.
- National Academy of Science, 1996: *National Science Education Standards*. <http://www.nap.edu/readingroom/books/nses/2.html#dsl> (Etat: 7.2.2008)
- Neeley, Stacia Dunn (2005): *Academic Literacy*. (Second Edition). New York: Pearson.
- Notter, Philipp & Arnold, Claudia (2003): Le passage aux études supérieures. Rapport relatif à un projet de la Conférence des directeurs de Gymnases suisses (CDGS) et de la Conférence des recteurs des universités suisses (CRUS). Berne: Secrétariat d'Etat à l'éducation et à la recherche SER (anc.: *Office fédéral de l'éducation et de la science OFES*) (= Dossiers OFES 2003, 5f).
- Nussbaumer, Markus & Sieber, Peter (1994): Texte analysieren mit dem Zürcher Textanalyseraster. Dans: Peter Sieber (éd.), *Sprachfähigkeiten - besser als ihr Ruf und nötiger denn je! Ergebnisse und Folgerungen aus einem Forschungsprojekt*. Aarau, Francfort-sur-le-Main, Salzbourg: Sauerländer, p. 142-185.
- Oelkers, Jürgen (1997): Geschichte und Nutzen der Projektmethode. Dans: Dagmar Hänsel (éd.), *Handbuch Projektunterricht* (p. 13-30). Weinheim, Bâle: Beltz.
- Oelkers, Jürgen (2007): *Die Qualität der Schweizer Gymnasien*. Berne: hep.
- Pagnossin, Elisabetta; Alliata, Roberta & Dozio, Edo (2005): Le travail de maturité. Dans: Erich Ramseier et al., (éd.) *Evaluation der Maturitätsreform 1995 (EVAMAR)*. Neue Fächerstruktur, Pädagogische Ziele, Schulentwicklung. Schlussbericht zur Phase 1. Berne: Secrétariat d'Etat à l'éducation et à la recherche SER, p. 171-191.
- Ramseier, Erich et al. (éd.) (2005): *Evaluation der Maturitätsreform 1995 (EVAMAR)*. Neue Fächerstrukturen, Pädagogische Ziele, Schulentwicklung. Schlussbericht zur Phase 1. Berne: Secrétariat d'Etat à l'éducation et à la recherche SER.
- Ramseier, Erich (2005): Stichprobenverfahren und Fehlerschätzung. *Evaluation der Maturitätsreform 1995 (EVAMAR)*. Ergänzungsband – Teil 1. Berne: Erziehungsdirektion des Kantons Bern.
- Ramseier, Erich et al. (non daté): *Befragung Abschlussklassen*. Berne: Conférence suisse des directeurs cantonaux de l'instruction publique et Office fédéral de l'éducation et de la recherche. http://www.sbf.admin.ch/evamar/FB/FB_EVAMAR_Abschluss_d.pdf (Etat: 11.4.2008)
- Ramseier, Erich; Allraum, Jürgen & Stalder, Ursula (2005): Neue Fächerstruktur und Ausbildungserfolg. Dans: Ramseier et al., *Evaluation der Maturitätsreform 1995 (EVAMAR)*. Neue

- Fächerstrukturen, Pädagogische Ziele, Schulentwicklung. Schlussbericht zur Phase 1. Berne: Secrétariat d'Etat à l'éducation et à la recherche SER, p. 45-155.
- Rheinberg, Falko (2002): Bezugsnormen und schulische Leistungsbeurteilung. Dans: Franz E. Weinert (éd.), *Leistungsmessungen in Schulen*. (2. Auflage). Weinheim, Bâle: Beltz, p. 59-71.
- Rychen, Dominique Simone & Salganik, Laura Hersh (2003): A holistic model of competence. Dans: Dominique Simone Rychen & Laura Hersh Salganik (éd.), *Key Competencies for a Successful Life and a Well-Functioning Society*. Gottingue: Hogrefe & Huber, p. 41-62.
- Schäfer-Koch, Karin (1997): Studienvorbereitung. Zur Genese eines Problems. Dans: Eckart Liebau et al. (éd.), *Das Gymnasium. Alltag, Reform, Geschichte, Theorie*. Weinheim, Munich: Juventa, p. 189-201.
- Schmid, Adrian (2002): Maturaarbeit: Last oder Lust. Dans: *Kollegiumsbote - Le Message du Collège*, 1, p. 10-11.
- Schmidt, Arno (1994): Das Gymnasium im Aufwind. Entwicklung, Struktur, Probleme seiner Oberstufe. Aachen-Hahn: Hahner.
- Schoch, Jürg (2003): Viel Aufwand für wenig Interdisziplinarität. Zur Reform der schweizerischen Gymnasien. Dans: *Neue Zürcher Zeitung*, 25.3.2003, p. 61.
- Schröder-Naef, Regula D. (1980): Von der Mittelschule zur Hochschule. Ergebnisse einer Befragung zur Reform der gymnasialen Oberstufe, zur Studienwahl und zu Übertritts- und Studienproblemen. Berne, Stuttgart: Haupt.
- Stocker, Eugen (2006): Regards sur le parcours gymnasial des jeunes Vaudois. Incidences des nouvelles réglementations sur la réussite des élèves. Lausanne: Unité de recherche pour le pilotage des systèmes pédagogiques (URSP).
- Tenorth, Heinz-Elmar (1994): *«Alle alles zu lehren». Möglichkeiten und Perspektive allgemeiner Bildung*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Trautwein, Ulrich, Lüdtke, Oliver & Husemann, Nicole (2006): Die Qualität der Studienvorbereitung in der gymnasialen Oberstufe: Eine Längsschnittstudie mit Absolventen von allgemeinbildenden und beruflichen Gymnasien. Dans: *Jahrbuch Jugendforschung*, 6, p. 47-67.
- Trautwein, Ulrich & Lüdtke, Oliver (2004): Aspekte von Wissenschaftspropädeutik und Studierfähigkeit. Dans: Olaf Köller, Rainer Watermann, Ulrich Trautwein & Oliver Lüdtke (éd.): *Wege zur Hochschulreife in Baden-Württemberg. TOSCA - Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien*. Opladen: Leske + Budrich, p. 327-366.
- Tschopp, Urs (2001): Zwischen Stühlen und Bänken - Lernen am Gymnasium zwischen Disziplin und Disziplinen. Versuch über die Interdisziplinarität und ihre Stellung in der gymnasialen Bildung. Dans: *Gymnasium Helveticum*, 5, p. 7-15.
- UCE [Union des conseils d'étudiants CH/FL] (2006): Schülerorganisationen: Endlich «zählt» die Maturaarbeit! *Communiqué de presse de l'UCE/USO du 27.12.2006*.
- Wallmeier, Martin (2007): Richtlinien zur Anfertigung wissenschaftlicher Arbeiten. Fribourg: Universität Fribourg, Departement für Betriebswirtschaftslehre, Lehrstuhl für Finanzmanagement und Rechnungswesen.
http://www.unifr.ch/finanzmanagement/lehre/Richtlinien_wiss_Arbeiten.pdf. (Etat: 15.9.2008)
- Weder, Hans (2005): Der Anspruch an die gymnasiale Bildung aus universitärer Sicht. Dans: *amv-aktuell*, Sonderheft 2005/1, p. 17-18.

- Weinert, Franz (2001): Concept of Competence: A Conceptual Clarification. Dans: Dominique Rychen & Laura Salganik (éd.), *Defining and Selecting Key Competencies*. Gottingue: Hogrefe & Huber, p. 45-65.
- Weinert, Franz (2002): Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. Dans: Franz Weinert (éd.), *Leistungsmessungen in Schulen* (2., unveränderte Aufl.). Gottingue: Hogrefe & Huber, p. 17-31.
- Wilhelmer, Hermann (2005): Fach(bereichs)arbeiten an berufsbildenden Schulen. Besichtigung eines Modells am Beispiel der HLW und Modeschule Klagenfurt. *Informationen zur Deutschdidaktik (ide)*, 3, p. 84-88.
- Winter, Felix (2006): Wie lernt man eine Maturaarbeit zu schreiben? Wissenschaftspropädeutisches Schreiben mit Diagnose und Rückmeldung. Dans: Otto Kruse, Katja Berger und Marianne Ulmi (éd.): *Prozessorientierte Schreibdidaktik. Schreibtraining für Schule, Studium und Beruf*. Berne, Stuttgart, Vienne: Haupt, p. 49-68.
- Wolter, Andrä (1997): Das deutsche Gymnasium zwischen Quantität und Qualität. Die Entwicklung des Gymnasiums und der Wandel gesellschaftlichen Wissens. Oldenburg: Bibliotheks- und Informationssystem der Universität Oldenburg (bis).
- Zillig, Elisabeth (2004): Gymnasiale Maturaarbeit und Interdisziplinäre Projektarbeit der Berufsmaturität. Ein Vergleich. Berne: Commission fédérale de la maturité professionnelle (CFMP/EMBK).

Tableaux et figures du chapitre VI

Liste des figures

Figure VI.1:	Genèse d'un travail de maturité	293
Figure VI.2:	Les conditions de réalisation des travaux de maturité dans le schéma de fonctionnement	298
Figure VI.3:	Les élèves dans le schéma de fonctionnement	298
Figure VI.4:	Séances spécifiques de préparation au travail de maturité dans les gymnases	311
Figure VI.5:	Modalités de choix du sujet (N=30)	312
Figure VI.6:	Consignes de forme (nombre de mentions)	313
Figure VI.7:	Variance du délai de préparation (mois)	314
Figure VI.8:	Pondération de la notation finale	316
Figure VI.9:	Satisfaction éprouvée à l'égard de l'encadrement fourni par les enseignants (en pour-cent)	324
Figure VI.10:	Fréquence des contacts avec la personne suivant le travail de maturité (en pour-cent)	325
Figure VI.11:	Fréquence des entretiens individuels de l'élève avec la personne suivant son travail de maturité (en pour-cent)	325
Figure VI.12:	Satisfaction des élèves à l'égard de la fréquence des entretiens (en pour-cent)	326
Figure VI.13:	Appréciation de la compétence de l'enseignant sur le fond (en pour-cent)	329
Figure VI.14:	Appréciation portée par les élèves sur l'utilité des séances et cours de préparation au travail de maturité (en pour-cent)	331
Figure VI.15:	Cadre institutionnel et individuel de préparation du travail de maturité	332
Figure VI.16:	Appréciation du volume de travail de préparation et de l'utilité (en pour-cent)	334
Figure VI.17:	Appréciation de l'utilité du travail de maturité dans l'itinéraire qualifiant et le développement personnel	335

Liste des tableaux

Tableau VI.1:	Hypothèses sur les aspects généraux de la qualité	300
Tableau VI.2:	Hypothèse relative au choix du sujet	300
Tableau VI.3:	Hypothèses relatives à la préparation	301
Tableau VI.4:	Hypothèses relatives aux conditions de réalisation	301
Tableau VI.5:	Hypothèses relatives au suivi des élèves	302
Tableau VI.6:	Hypothèses relatives à la motivation et à la perception de soi	303
Tableau VI.7:	Hypothèses relatives à la valeur intrinsèque et extrinsèque du travail	303
Tableau VI.8:	Population étudiée (élèves en première année de formation gymnasiale au cours de l'année scolaire 2004-2005)	304
Tableau VI.9:	Population et échantillon de gymnases	304
Tableau VI.10:	Objectifs, 2006 (N=30)	310

Tableau VI.11: Choix du sujet selon la discipline générale (pourcentages, réponses multiples admises)	320
Tableau VI.12: Choix du sujet du travail de maturité selon la région linguistique (pourcentages)	321
Tableau VI.13: Motifs de choix du sujet (pourcentages, réponses multiples admises)	321
Tableau VI.14: Notes données aux travaux de maturité (pourcentages)	322
Tableau VI.15: Choix de la personne encadrant le travail de maturité selon la région linguistique (pourcentages, réponses multiples admises)	323
Tableau VI.16: Aide demandée à l'enseignant (réponses multiples admises)	327
Tableau VI.17: Appréciation de l'assistance reçue (valeur moyenne, écart type)	328
Tableau VI.18: Autres sources d'aide (réponses multiples admises)	329
Tableau VI.19: Séances et cours préparatoires (réponses multiples admises)	331
Tableau VI.20: Appréciation du volume de travail individuel effectif	333
Tableau VI.21: Apports du travail de maturité	336
Tableau VI.22: Echelles «organisation du travail» et «compétences cognitives»	337
Tableau VI.23: Corrélation entre l'utilité et les apports perçus	337
Tableau VI.24: Proportion d'écarts d'évaluation ≤ 2 points de l'échelle (pourcentage par critère)	339
Tableau VI.25: Répartition des travaux selon la discipline	340
Tableau VI.26: Saturation factorielle des composantes de la qualité des travaux de maturité	341
Tableau VI.27: Echelle «qualité du fond»	342
Tableau VI.28: Echelle «qualité de la langue»	343
Tableau VI.29: Echelle «qualité de la forme»	343
Tableau VI.30: Corrélations entre les trois échelles (qualité du fond, de la langue et de la forme)	343
Tableau VI.31: Valeurs moyennes et écarts types des critères de l'échelle «qualité du fond»	344
Tableau VI.32: Valeurs moyennes et écarts types des critères de l'échelle «qualité de la langue»	345
Tableau VI.33: Valeurs moyennes et écarts types de l'échelle «qualité de la forme»	346
Tableau VI.34: Corrélation entre l'appréciation globale et les échelles de qualité du fond, de la langue et de la forme	347
Tableau VI.35: Corrélation entre l'appréciation portée sur l'assistance reçue de l'enseignant suivant le travail et les échelles de qualité	348
Tableau VI.36: Corrélation entre l'utilité perçue et les trois échelles d'appréciation de la qualité du fond, de la langue et de la forme	350

Annexe du chapitre VI

Annexe VI.A: Vue d'ensemble des hypothèses testées

N°	Hypothèse	Confirmée	
		oui	non
1	Aspects généraux de la qualité		
1a	Il existe une corrélation positive entre la qualité du contenu, celle de la forme et celle de la langue des travaux de maturité.	x	
1b	Il existe des écarts significatifs liés au sexe, à la région linguistique et à la forme de travail dans la qualité du contenu, de la forme et de la langue.		
	1b1: sexe		x
	1b2: région linguistique	x	
	1b3: forme de travail	x	
1c	La discipline dont relève un travail de maturité n'a pas d'influence significative sur les divers aspects de sa qualité.	(x)	(x)
	1c1: La discipline dont relève un travail de maturité n'a pas d'influence significative sur la qualité de son contenu.	x	
	1c2: La discipline dont relève un travail de maturité n'a pas d'influence significative sur sa qualité formelle.		x
	1c3: La discipline dont relève un travail de maturité n'a pas d'influence significative sur sa qualité linguistique.		x
1d	Les travaux interdisciplinaires ne présentent pas de différence de qualité significative avec les travaux relevant clairement d'une discipline.	x	
2a	Choix du sujet		
	Plus l'élève est libre dans le choix de son sujet, meilleure devient la qualité du contenu du travail de maturité.		x
2b	Préparation		
2b1	Une préparation intense (semaine de travail intensif ou sur un projet, travail par projet, initiations, etc.) a un effet positif sur la qualité du travail.	(x)	(x)
	2b1.1: sur la qualité du contenu	x	
	2b1.2: sur la qualité formelle		x
	2b1.3: sur la qualité linguistique	x	
2b2	Les travaux de maturité des élèves qui n'ont assisté qu'à des séances d'information ou n'ont même suivi aucune préparation ont tendance à présenter une qualité inférieure.	x	
2b3	Les séances de préparation ressenties comme utiles ont un effet positif sur la qualité du travail.	x	
2c	Conditions de réalisation		
2c1	Il existe un lien entre le délai de préparation imparti par l'établissement et la qualité du travail de maturité.		x
2c2	L'existence et la connaissance d'un guide se traduisent par une meilleure qualité des travaux de maturité.		x

N°	Hypothèse	Confirmée	
		oui	non
2c3	La transparence perçue des modes opératoires a un effet positif sur la qualité des travaux de maturité.		x
2c4	La transparence perçue des critères d'évaluation a un effet positif sur la qualité des travaux de maturité.		x
2c5	Il n'existe aucune corrélation entre l'appréciation portée sur les critères d'évaluation et la qualité.		x
3 Suivi des élèves			
3a	Mieux l'élève est suivi, meilleure est la qualité du travail de maturité.		x
3b	La satisfaction éprouvée à l'égard de l'assistance reçue dans la formulation de la problématique et sur le contenu du travail présente une corrélation positive avec la qualité du contenu du travail de maturité.		x
3c	La satisfaction éprouvée à l'égard de l'assistance reçue sur les bases formelles, la structuration et l'articulation du travail présente une corrélation positive avec la qualité formelle du travail de maturité.	(x)	(x)
	3c1: bases formelles	x	
	3c2: structuration et articulation		x
3d	Plus l'élève est libre dans le choix de l'enseignant qui l'encadrera, mieux le suivi est ensuite apprécié.	x	
3e	Il n'existe pas de lien entre la fréquence et la forme des contacts d'encadrement d'une part, et la qualité du travail de maturité de l'autre.	x	
4a Motivation et perception de soi			
4a1	Il existe une corrélation positive entre le temps que l'élève rapporte avoir consacré à son travail de maturité et la qualité de ce dernier.		x
4a2	Il existe une corrélation positive entre l'assistance reçue de l'entourage personnel (parents, frères et sœurs, camarades) et la qualité du travail de maturité.	(x)	(x)
	4a2.1: parents		x
	4a2.2: frères et sœurs	x	
	4a2.3: camarades	x	
4a3	La connaissance préalable du sujet a un impact positif sur la qualité de fond du travail de maturité.		x
4a4	Les compétences (connaissances préalables et autonomie) que pense posséder l'élève présentent une corrélation positive avec la qualité du travail.		x
4a5	L'intérêt déclaré <i>a posteriori</i> par l'élève présente une corrélation positive avec la qualité.		x
4b Valeur intrinsèque et extrinsèque du travail			
4b1	Il existe une corrélation positive entre la qualité du travail et l'appréciation de son utilité dans l'obtention de futures qualifications.		x
4b2	Il existe une corrélation positive entre l'appréciation par l'élève de l'utilité du travail et celle de l'accroissement de ses propres compétences.		x
4b3	Il existe une corrélation positive entre l'appréciation du travail par l'enseignant et l'appréciation par l'élève des acquis ainsi obtenus.	x	

Annexe VI.B: Grille d'évaluation des conditions cadre du travail de maturité

Exemple: Formulaire d'évaluation de l'établissement X

Critère	Prédicat/remarques	
Règles écrites internes	oui	Règlement, critères de réussite, directives
Compétence organisationnelle	oui	Service de coordination (comprenant un membre de la direction)
Guide	oui	Dans les directives Thèmes: bases légales, objectifs, organisation du travail (seul/en groupe), choix du sujet, suivi, volume, déclaration sur l'honneur, présentation, évaluation, voies de recours, délais, organisation
Définition des objectifs	oui	Dans les directives Thèmes: travail autonome, forme écrite, méthodes adaptées à la discipline; présentation orale
Préparation	oui	Séance d'information
Choix du thème	libre	Dans une discipline précise ou interdisciplinaire
Types		(1) écrit, (2) écrit avec commentaire oral, (3) pratique/plastique
Travail individuel/en groupe		Pas de conditions précisées
Processus de travail documenté	oui	Journal de bord
Instructions sur le temps à investir	non	
Instructions sur la forme	oui	Longueur: 15 à 20 pages sans annexe; format recommandé
		Structure Titre/Table des matières/Préface – Introduction – Partie principale – Résumé – Bibliographie
		Conventions de citation
Contrat de travail	non	
Durée d'enseignement	oui	2 HHS
Préparation des enseignants	non	Pas d'indications sur la préparation et les év. compensations prévues
Nombre de travaux à suivre	2/4	Au minimum/au maximum
Temps à consacrer par travail suivi	n. i.	L'enseignant doit mener des entretiens réguliers.
Charge d'établir la relation de suivi	E	(E = élève, P = enseignant)
Personnes habilitées au suivi		Maîtres et chargés de cours (internes)
Conséquences	ND, C, AP	(ND = note de discipline, C = certificat, AP = accès à la présentation)
Présentation (durée, examinateurs)		15 min. prés. /15 min. discussion, responsable du suivi + collègue
Pondération (travail/présentation)	66/33	Travail écrit: évalué sur le contenu, la forme et la démarche
Appréciation du travail		Verbale et écrite (note)
Evaluation de groupe		Note de groupe; note individuelle possible en accord avec les élèves
Temps imparti aux élèves		52 semaines (du choix définitif du sujet à la remise du travail)

Annexe VI.C: Facteurs de pondération

(1) Pondération par région linguistique

Région linguistique	Facteur de pondération
Suisse alémanique	0.876325
Suisse romande	1.2
Tessin	1.170909

(2) Pondération par sexe

Sexe	Facteur de pondération
féminin	0.9533185
masculin	1.0670425

Remarque: Les facteurs de pondération ci-dessus ont été employés pour augmenter la comparabilité avec les résultats d'EVAMAR I. En dehors de cette comparaison, les résultats décrits dans le présent chapitre se fondent exclusivement sur des données non pondérées.

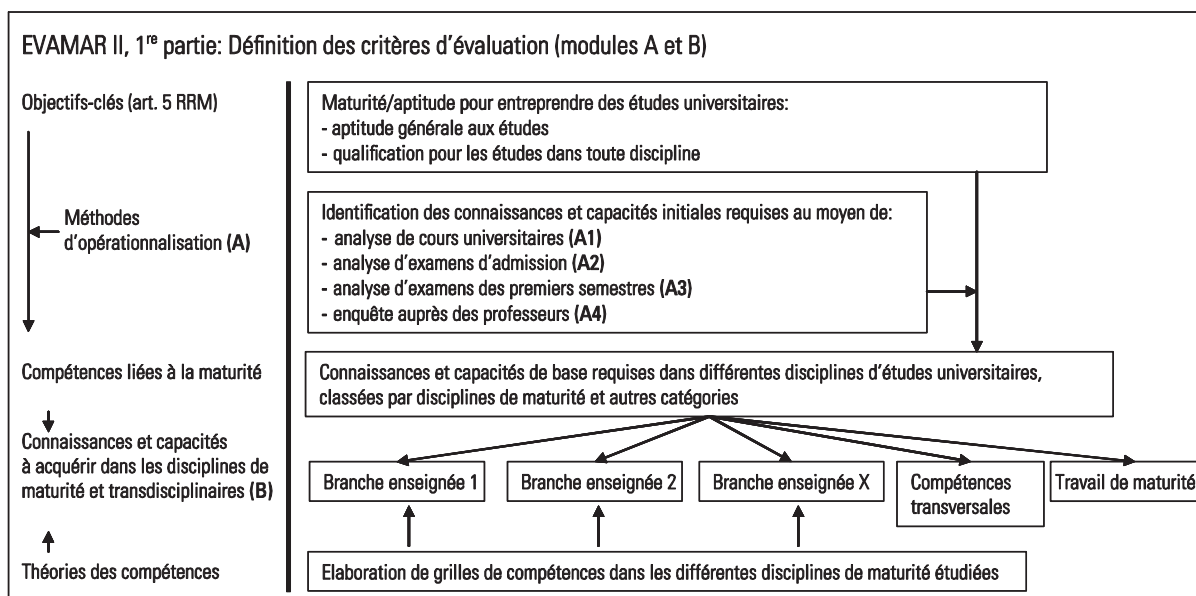
VI Résumé

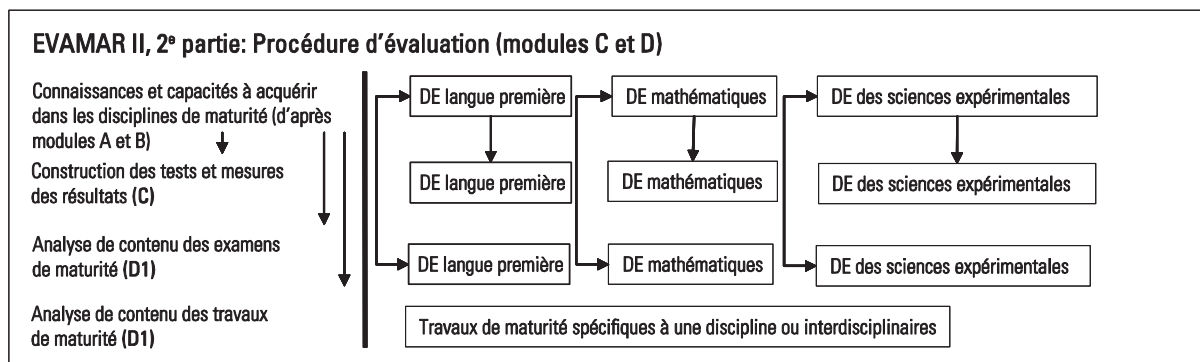
1 Objets de l'étude

Un relevé évaluatif du niveau de formation doit logiquement se référer aux objectifs de cette formation. Selon l'art. 5 du RRM 95, la formation gymnasiale doit premièrement permettre d'acquérir la «maturité requise pour entreprendre des études supérieures» (maturité nécessaire ou aptitude aux études supérieures) et, deuxièmement, préparer les élèves à «assumer des responsabilités au sein de la société actuelle» (maturité sociale étoffée d'une vaste culture générale). Les incidences de ces deux objectifs sur les programmes sont parfois importantes et vont tantôt dans le même sens, tantôt dans des sens divergents. Il serait extrêmement coûteux de transposer entièrement ces objectifs en compétences concrètes que les gymnasiennes et gymnasiens sont censés acquérir dans toutes les disciplines et tous les domaines interdisciplinaires et de mesurer systématiquement le niveau réel de chacun, comme cela était prévu dans le premier projet proposé par le mandataire. C'est la raison pour laquelle le mandant a choisi de limiter les enquêtes réalisées dans le cadre d'EVAMAR II aux objets ou sous-projets esquissés dans la figure 1, en concrétisant certains éléments seulement de l'aptitude aux études supérieures, qui est l'objectif principal de la formation gymnasiale. Le second objectif (maturité sociale étoffée d'une vaste culture générale) n'a de ce fait été inclus que partiellement (aspects communs seulement) dans la définition opérationnelle des objectifs choisie par le mandant.

Seuls *certain*s aspects de l'aptitude aux études supérieures peuvent donc être commentés à la lumière des enquêtes retenues pour EVAMAR II. Le second objectif, préparer les gymnasiennes et gymnasiens à «assumer des responsabilités au sein de la société actuelle», n'a pas été étudié en tant que tel (cf. les limites de l'étude décrites au point 7). Il faudra tenir compte de ces restrictions pour interpréter les résultats et en déduire des mesures concernant l'organisation future du gymnase.

Figure VII.1: Enquêtes retenues pour EVAMAR II





Légende:

- Les lettres A à D désignent les différents modules du projet.

- DE: abréviation de «discipline enseignée»

Le présent résumé s'étendra davantage sur les modules C et D que sur les modules A et B (pré-analyses) afin de se concentrer sur les principaux résultats de l'enquête.

2 Résultats I: Identification des performances à mesurer (pré-analyse)

Analyse du matériel de cours et des examens

Une analyse du matériel d'enseignement et des examens a été menée dans les sous-projets A1 et A3, afin de déterminer les connaissances et compétences requises pour suivre les deux premiers semestres universitaires. L'étude couvrait les 16 branches les plus choisies durant l'année académique 2004/2005, étudiées par environ 70% des personnes immatriculées en Suisse. Pour chaque branche, nous nous sommes adressés aux deux ou trois universités qui regroupent ensemble plus de la moitié des étudiants ayant choisi cette orientation. Les réponses et documents reçus couvraient les 16 branches retenues, bien que tous les professeurs contactés n'aient pas fourni leurs supports de cours et leurs examens. S'agissant de l'analyse des premiers examens menée dans le cadre du sous-projet A3, il faut préciser en outre que nous n'avons pas pu disposer d'examens pour chaque branche, vu qu'un certain nombre de filières d'études ne comportent pas d'examen au cours des deux premiers semestres, sur lesquels portait notre analyse.

L'analyse du matériel d'enseignement et des examens portait sur les relations entre disciplines de maturité et branches d'études (figure 2). Elle a permis de constituer une collection d'éléments de connaissances et compétences classés dans une base de données, qui est donc plus que le simple reflet du Plan d'études cadre et des programmes des écoles de maturité. Il a fallu concevoir intégralement la méthodologie d'analyse et de codification appliquée dans le cadre du module A (et B) d'EVAMAR II.

Figure VII.2: Relation entre disciplines de maturité et branches d'études

Connaissances et capacités en...	Mathématiques	Langue première	Géographie	Physique	Biologie	Discipline Y
... pour la discipline d'études						
Germanistique	GM	GL1	GG	GP	GB	GY
Droit	DM	DL1	DG	DP	DB	DY
Sociologie	SM	SL1	SG	SP	SB	SY
Biologie	BM	BL1	BP	BP	BB	BY
X	XM	XL1	XG	XP	XB	XY
	↓	↓	↓	↓	↓	↓
Compétences d'études à développer en...	Mathématiques	Langue première	Géographie	Physique	Biologie	Discipline Y ou compétences transversales
	↑	↑			↑	
Modèles de compétences élaborés pour les mathématiques, la langue première et la biologie						

Légende: GM = connaissances et capacités à acquérir au gymnase en mathématiques (M) pour étudier la germanistique (G) à l'université, etc.

A partir des cours et examens analysés pour l'ensemble des branches d'études, on peut dégager globalement les tendances suivantes à propos des connaissances et compétences supposées acquises et mobilisables dans les différentes disciplines de maturité:

- Sur le plan de la catégorisation cognitive, on trouve nettement plus d'éléments du type connaissances factuelles que de connaissances procédurales ou métacognitives, qui sont assez rarement supposées acquises. Cette fréquence s'observe en particulier pour les éléments relevant des langues enseignées à l'école et concerne avant tout les trois langues premières (allemand, français, italien) et l'anglais. Beaucoup d'entre eux sont toutefois également donnés par le latin et le grec, une grande partie relevant simultanément de la langue première. La mise en relation des langues anciennes avec la maîtrise des concepts abstraits réapparaît dans bon nombre de branches scientifiques.
- Sur le plan des disciplines scolaires, la catégorisation est très variable et, comme on pouvait s'y attendre, dépend étroitement de la branche analysée. Dans l'analyse de la biologie universitaire, par exemple, on trouve fréquemment des éléments relevant de la biologie et de la chimie enseignées à l'école; ou encore en sciences économiques, on constate que l'enseignement universitaire présuppose un éventail de notions économiques déjà étonnamment étendu.
- Les connaissances de grec et de latin peuvent le plus souvent passer pour utiles, mais ne sont pas nécessairement considérées comme indispensables, par exemple pour les études en droit, en histoire ou en pédagogie; une connaissance approfondie du latin et du grec pourrait néanmoins, on l'a vu, présenter généralement des avantages surtout pour le vocabulaire, qui est lui-même très important pour l'ensemble des branches d'études. Mais il faut tenir compte du fait que les termes techniques peuvent être compris ou acquis même sans études linguistiques de ce type.
- Outre de vastes connaissances de la langue première et un riche vocabulaire savant, on exige dans bien des cas des étudiants une bonne compréhension de l'écrit, en sciences humaines et sociales notamment, où l'histoire, la sociologie et la pédagogie font fréquemment mention d'éléments ayant trait à l'étude de textes complexes et requérant un niveau adéquat de «littératie».
- S'agissant des connaissances en sciences expérimentales, on ne constate pratiquement aucune différence quant au nombre d'éléments de physique et de chimie requis par chacune des branches non expérimentales (c'est-à-dire, dans notre cas, toutes les branches sauf la physique, la biologie et la médecine). En revanche, des notions de biologie sont nettement plus souvent exigées dans les sciences humaines et sociales que des concepts de physique ou de chimie.

- Des acquis sont également attendus dans le domaine de compétence transversale que constitue le travail scientifique. Mais certaines branches proposent également des cours ou séminaires d'introduction au travail scientifique.
- L'analyse des connaissances requises dans le cadre des exercices et examens montre que, le plus souvent, ces derniers font peu référence à des éléments devant déjà avoir été acquis au gymnase; parmi les compétences cognitives requises lors des examens, on trouve en majorité des types de compétences comme «mémoriser les connaissances», «comprendre», «appliquer» et «analyser». Par contre, on exige assez peu des étudiants qu'ils sachent «évaluer, créer» ou «générer, planifier», du moins pendant les deux premiers semestres.

Les résultats obtenus dans les sous-projets A1 et A3 devaient essentiellement fournir une base précieuse aux tests réalisés dans la suite du projet pour la langue première, les mathématiques et la biologie enseignées à l'école.

Enquête auprès du corps professoral

L'objectif de l'enquête auprès des professeurs était de savoir quelles connaissances et compétences spécifiques aux disciplines et transversales ils supposaient acquises chez leurs étudiants des premier et deuxième semestres. Nous leur avons également demandé s'ils constataient des lacunes dans le bagage des nouveaux étudiants. Un questionnaire a été adressé à 320 professeurs d'universités alémaniques, dont les cours avaient précédemment aussi été retenus pour l'analyse du matériel d'enseignement. Le taux de réponse fut d'environ 44%. Nous en résumons dans ce qui suit les principaux constats, jugés dans l'ensemble également applicables à leur région linguistique par quelques professeurs de Suisse romande et du Tessin.

Nous avons tout d'abord prié les professeurs d'indiquer quelle importance ils attachaient aux connaissances et compétences spécifiques acquises dans les différentes disciplines gymnasiales et en informatique (utilisation des TIC et science informatique) pour la compréhension et une fréquentation utile de leurs cours. Si l'on classe sur une échelle de 1 à 4 les moyennes des réponses obtenues, on constate que les matières jugées les plus importantes sont, dans l'ordre, la langue première, l'anglais, l'informatique (utilisation des TIC) et les mathématiques. La biologie et la physique se classent au milieu de l'échelle, et la chimie en haut du tiers inférieur.

Si l'on classe les réponses à cette question par branche d'études, la langue première conserve dans presque tous les cas le premier rang, et l'anglais et l'informatique (utilisation des TIC) restent dans le haut du tableau. Pour les autres disciplines, on obtient des différences de classement parfois considérables. En sciences sociales et dans les branches linguistiques, les mathématiques sont sans surprise considérées comme assez peu importantes, mais elles le restent pour bien d'autres branches. La priorité va naturellement à la discipline de maturité correspondant à la branche étudiée.

Les professeurs ont ensuite été invités à indiquer concrètement quelles connaissances et compétences ils supposaient acquises dans les différentes disciplines. Les résultats ont montré qu'il pouvait y avoir des divergences quant à la manière de concevoir le programme de l'une ou l'autre discipline. Lorsque les professeurs parlent de la grande importance de la langue première, ils ont surtout à l'esprit la «maîtrise de la langue», la grammaire, la syntaxe, etc. Seuls certains germanistes et un angliciste ont exprimé des attentes correspondant aux véritables contenus centraux de la germanistique que sont l'histoire de la langue ou de la littérature, la méthodologie interprétative et herméneutique, les éléments fondamentaux de la linguistique, de la dramaturgie, etc. La notion de la discipline «langue première» recouvre donc des représentations variables du contenu enseigné. Si la langue première est généralement jugée comme la discipline de maturité la plus importante, c'est essentiellement pour les compétences linguistiques qu'elle procure. Il en va de même pour l'anglais, où l'on pense d'abord à la capacité de traiter la littérature spécialisée en anglais (et non par exemple à un bagage en histoire de la littérature).

La question suivante portait sur les lacunes constatées chez les nouveaux étudiants parmi les connaissances et compétences jugées importantes en réponse aux questions précédentes. Des lacunes ont été signalées dans presque toutes les disciplines, particulièrement dans la langue première et en mathématiques. Sur le plan de la langue première, il s'agissait à nouveau pour l'essentiel d'éléments de «maîtrise de la langue», dont la «pauvreté» a été critiquée surtout par les professeurs non germanistes.

Les professeurs avaient ensuite à évaluer sur une échelle de 7 l'importance des compétences transversales pour la compréhension et une fréquentation utile de leurs cours. Toutes les compétences indiquées ont été jugées de très importantes (proche de 7) à au moins moyennement importantes (un peu plus de 4), dans l'ordre suivant: (1) apprendre de manière autonome, (2) assumer ses responsabilités dans son apprentissage et ses travaux personnels, (3) travailler de manière autonome, (4) faire preuve d'esprit critique, (5) être capable de résoudre les problèmes, (6) savoir s'exprimer par écrit, (7) organiser son temps avec efficacité et définir des priorités, (8) gérer la pression, (9) comprendre l'oral dans le cadre des cours ex cathedra, (10) savoir s'exprimer oralement, (11) préparer des examens sur une vaste matière, (12) développer des idées nouvelles, (13) enregistrer beaucoup de matière en peu de temps, (14) faire une recherche, (15) travailler en équipe, (16) parler en public. Les professeurs interrogés avaient en outre la possibilité d'indiquer dans lesquelles de ces compétences ils constataient des lacunes. Les réponses les plus fréquentes furent les suivantes: expression écrite (42,4%), esprit critique (35%), travail autonome (30%), apprentissage autonome (26%), expression orale (21%), responsabilité dans son apprentissage et ses travaux personnels (20,6%). Les moins fréquentes étaient: se présenter devant un public (3%) et travailler en équipe (3%).

Les résultats de l'enquête n'ont pratiquement pas imposé de correction dans les conclusions des analyses A1 et A3 en vue des tests, mais ils n'en sont pas moins intéressants en tant que constats particuliers. Quelques exceptions sont toutefois à signaler: pour la langue première, nous avons rajouté au test un volet «compétences en grammaire et en orthographe»; nous avons conçu un questionnaire complémentaire qui invite à faire une autoévaluation des principales compétences transversales; enfin, les résultats de cette enquête ont corroboré le choix de la biologie parmi les sciences expérimentales (cf. plus loin).

Analyse des examens d'admission pratiqués actuellement (A2)

Ce sous-projet visait à obtenir des indications complémentaires sur les compétences qu'est censé posséder quiconque entreprend des études académiques, en analysant les examens d'entrée pratiqués par les universités suisses. Ce sont avant tout des examens d'admission proposés par la plupart des universités comme possibilité de faire valoir ses qualifications si l'on n'est pas titulaire d'une maturité suisse. Un premier survol montre que ces examens se réfèrent en règle générale au programme des examens fédéraux de maturité et qu'ils sont donc très semblables aux examens passés dans le cadre du gymnase. Ces derniers étant l'objet de l'analyse réalisée au sous-projet D1, cette catégorie d'examen d'admission a donc été écartée de la suite de l'étude pour parer au risque d'établir des clés d'interprétation circulaires. La seconde catégorie d'examen d'entrée est formée des tests d'aptitude à proprement parler, s'appuyant sur des méthodes de prédiction de la réussite dans les études qui soient expressément formulées, dotées d'une assise théorique et, si possible, validées aussi empiriquement. Il s'avère que le seul instrument répondant en Suisse à ces critères est le test d'aptitude aux études de médecine. Utilisé par différentes universités pour sélectionner les candidates et candidats aux études de médecine en raison du nombre de places limité, il est constitué de questions sollicitant des compétences transversales (ce n'est pas un test de connaissances). Puisque l'existence de ces questions offrait la possibilité de recenser davantage de compétences ayant un rôle à jouer dans les études, et donc de compléter les tests spécifiques par discipline prévus dans EVAMAR II, le Groupe de projet EVAMAR II a décidé de présenter également des questions tirées d'éditions précédentes de ces tests d'aptitude, dans le cadre du module C réalisé en Suisse alémanique et ce, bien que ce volet n'eût pas été prévu dans le projet de mandat.

3 Modélisation des compétences (B)

La planification d'EVAMAR II prévoyait initialement de condenser directement en modèles de compétence les résultats des analyses des sous-projets A1 à A4. La réalisation de ce projet n'a toutefois pas tardé à se révéler trop difficile pour diverses raisons, la principale étant que les textes de cours et exercices analysés dans le module A se prêtaient certes à établir des relations directes avec les connaissances et compétences absolument indispensables à une appréhension aisée, mais ne fournissaient pas assez d'informations susceptibles de composer un modèle de compétence cohérent, correspondant à l'état actuel des connaissances théoriques en matière d'apprentissage et validé empiriquement, selon la conception prônée actuellement dans les discussions théoriques autour des standards de formation. De surcroît, les unités de sens ou de connaissance et les capacités cognitives ainsi établies ne recouvraient pas totalement les exigences des universités. Il s'agissait seulement d'un échantillon, lui-même encore trop étroit pour servir de base unique à un modèle de compétence embrassant l'ensemble de l'aptitude aux études supérieures. Il s'est donc avéré nécessaire d'analyser à la lumière des objectifs d'EVAMAR II la pertinence des modèles de compétence qui existaient dans chaque discipline, de les reprendre ou de les perfectionner. Ces adaptations dûment documentées ont essentiellement servi à préciser, à l'aide des connaissances et compétences spécifiques identifiées dans le projet A, les descriptions plutôt générales et formelles des domaines couverts par les modèles de compétence, dans l'optique de l'interface qu'étudie EVAMAR II. Des modèles ont donc été développés pour les trois disciplines étudiées: la langue première, les mathématiques et la biologie. Pour respecter la définition de la notion, ou plutôt du concept construit «modèle de compétence» attesté dans les discussions actuelles et qui a une vocation générale, nous avons appelé «grilles de compétence» les résultats de notre modélisation.

Nous avons obtenu de cette manière, pour les trois disciplines, des champs de compétence s'accordant avec les connaissances et compétences identifiées dans le projet A. Le module C a ensuite construit pour ces champs des items se rapportant aux contenus spécifiques aux disciplines identifiées dans le projet A. Tous les items du test peuvent donc se référer aussi bien à la modélisation choisie qu'aux supports de cours universitaires. Ce «tramage des points d'intersection» concret et systématique opéré par le test ou, plus exactement, par la mesure des performances effectuée dans le projet EVAMAR II n'avait encore jamais été réalisé dans un contexte semblable, c'est donc une première. Voici les modèles de compétence qui ont été jugés les plus dignes d'intérêt pour EVAMAR II: le cadre européen commun de référence pour les langues (CECRL), la modélisation des compétences en mathématiques de la troisième étude internationale sur les mathématiques et les sciences (TIMSS III) et les exigences communes pour l'examen de baccalauréat allemand en biologie émises par la Conférence permanente des ministres de l'éducation et des affaires culturelles des länder (KMK) mises en relation avec l'articulation des chapitres des quatre ouvrages alémaniques standard pour l'enseignement de la biologie au niveau gymnasial (secondaire II). On notera que, du fait de l'orientation du projet, les grilles de compétence ne correspondent pas nécessairement aux programmes des trois disciplines. Tel est notamment le cas pour la langue première. La grille de compétences et le test qui s'y réfère omettent en effet un grand nombre d'éléments propres au cursus de cette discipline, dont l'importance réside dans d'autres raisons que le fait de préparer à une multitude de branches d'études.

4 Résultats II: Mesure des performances (C)

Description des tests réalisés et de l'échantillon

Le module C a été le cadre d'enquêtes auprès des bacheliers et bachelères. Elles se composaient de tests dans les trois disciplines de la langue première, des mathématiques et de la biologie, d'un questionnaire complémentaire et d'un test des compétences transversales (TCT). Ces deux derniers éléments de l'enquête se sont révélés nécessaires en raison des pré-analyses décrites ci-dessus. C'est le mandant qui a souhaité appliquer l'enquête à une discipline scientifique en lieu et place de l'anglais initialement proposé. Le choix concret entre la

biologie, la chimie et la physique n'avait pas encore été fait au moment de l'attribution du mandat. Le groupe de projet a opté pour la biologie surtout parce que cette discipline renferme des éléments utiles à un plus grand nombre de branches d'études que les deux autres (cf. les résultats de l'enquête auprès des professeurs) et parce que le type de performances que l'on y teste présente davantage de différences avec les éléments déjà mesurés en mathématiques et dans le TCT. Le questionnaire demandait notamment, en complément à EVAMAR I, d'auto-évaluer d'autres compétences faisant partie de l'aptitude aux études supérieures. Il a été rempli de manière encore plus complète en Suisse alémanique, en raison du temps imparti. Le TCT contient quant à lui des items ayant des corrélations, bien attestées par d'autres enquêtes, avec la réussite des études de sciences. Ses items sont validés et proviennent des quatre volets du test d'aptitude aux études de médecine qui ont une valeur prédictive particulièrement haute pour les examens dans les branches d'études universitaires. Ce test additionnel n'a été réalisé qu'en Suisse alémanique car la direction du projet a estimé que l'on ne pouvait plus exiger davantage de temps pour ces tests que ce qui avait été planifié à l'origine en Suisse romande et au Tessin.

La construction des items du test s'est basée sur les grilles de compétence établies dans le module B, sur les éléments de connaissance et de compétence saisis dans la base de données résultant des sous-projets A1 et A3 et sur le matériel de cours concret en rapport avec le sujet. Nous avons veillé à obtenir une répartition aussi représentative que possible entre les champs de compétence d'une part et entre les éléments de connaissance et de compétence identifiés d'autre part. A l'exception, déjà mentionnée, du domaine de la langue première, il n'a pas été nécessaire de corriger cette base à la lumière des résultats de l'enquête auprès du corps professoral réalisée dans le sous-projet A4. Afin de pouvoir intégrer toute l'étendue des grilles de compétence, nous avons conçu une somme d'items dont la résolution complète aurait pris à l'élève 720 minutes. Le temps imparti à l'enquête n'était pas suffisant pour soumettre à chacun la totalité des items. Nous disposions seulement de 45 minutes par test. Pour garantir une évaluation fiable des aptitudes des bacheliers et bachelères, mais aussi de la difficulté des items, nous avons choisi une configuration tournante dans laquelle les items étaient reliés entre eux (une matrice multicritères reliant des blocs d'items similaires de manière que les connaissances et compétences relevées chez la personne ne se rapportent pas uniquement au cahier de test qu'elle a rempli). Les cahiers ont été distribués au hasard aux élèves, en veillant à ce que chaque item soit traité dans chaque classe par un nombre semblable de personnes. Les critères déterminants lors du choix des types d'items étaient, à l'instar d'enquêtes de taille comparable (TIMSS, PISA), ceux de l'objectivité et de l'économicité de l'analyse. Nous avons retenu le moins possible d'items à réponse ouverte et circonstanciée (la plupart en langue première) car ils auraient exigé de lourds travaux de codification pour assurer l'objectivité de la correction. En nous limitant en outre à la forme écrite, nous avons restreint le nombre de champs de compétence testables pour la langue première (par ex. pas de compétences orales ou auditives).

La constellation d'instruments choisie permettait globalement de mesurer de multiples aspects à l'aune de l'aptitude aux études supérieures (connaissances spécifiques aux disciplines, compétences linguistiques générales, compétences en mathématiques, en biologie et plus généralement en sciences).

L'enquête a été réalisée à plus ou moins la même période dans tous les établissements. La condition était que les gymnasiennes et gymnasiens n'aient plus que trois semaines de cours au maximum avant les examens de maturité. En raison des différences de calendrier d'un établissement à l'autre, la période de l'enquête s'est donc étalée de fin avril à début juillet 2007.

La population de base comprend les bacheliers et bachelères ayant passé leur maturité en Suisse en été 2007, à l'exclusion de ceux de Bâle-Campagne et de Genève. A Bâle-Campagne, en effet, les examens de maturité n'ont pas lieu avant l'automne, ce qui aurait empêché toute comparaison. Et à Genève, les lycées se sont opposés dès le départ à l'enquête, la plupart des établissements refusant fin 2006 de livrer les indications qui auraient permis de définir l'univers statistique. Le groupe de pilotage du projet EVAMAR II n'a pas voulu, pour diverses raisons, obtenir une participation imposée par la voie hiérarchique. Une participation forcée aurait sans doute entraîné un fort risque de falsification nuisible à la validité des résultats. L'échantillon a été stratifié afin de permettre différentes comparaisons entre sous-populations. Toutes les classes invitées n'ont pas participé à l'enquête. Le seul canton où le taux de participation s'est révélé problématique car extrêmement bas pour divers

motifs est le Tessin (cf. pour les détails le rapport principal), raison pour laquelle les résultats tessinois obtenus lors de ce sous-projet n'ont pu être inclus dans les analyses. Le taux de réponse (hors Tessin) est de 91% pour les classes et de 85% pour les personnes, cet écart s'expliquant par les absences enregistrées le jour de l'enquête. Nous avons donc eu au total des données analysables pour 3800 personnes.

Résultats généraux

Un peu plus de la moitié des items en moyenne ont été résolus correctement (nettement plus en «compétence grammaticale et orthographe»). Sur l'ensemble des domaines de compétence, les aptitudes personnelles des bacheliers et des bachelères correspondent à un niveau de difficulté moyen. La dispersion entre les individus est considérable. Si l'on excepte le domaine «compétence grammaticale et orthographe», le nombre maximal de points possibles a chaque fois été atteint. En poussant l'analyse aux valeurs agrégées sur l'ensemble des classes, on obtient également une dispersion considérable entre les moyennes par classe. La meilleure classe a en moyenne résolu correctement deux fois plus d'items que la moins bonne.

En mathématiques, les personnes testées ont en moyenne résolu correctement un peu moins de la moitié des items. Sur l'ensemble des domaines de compétence, les aptitudes personnelles des bacheliers et bachelères correspondent, en moyenne, à un niveau de difficulté très légèrement inférieur à la moyenne. Ici aussi, la dispersion des individus est considérable: le nombre maximal de points a été atteint dans tous les domaines. En poussant l'analyse aux valeurs agrégées sur l'ensemble des classes, on constate également une dispersion importante entre les moyennes par classes. En moyenne, la meilleure classe a résolu correctement presque trois fois plus d'items que la moins bonne.

Etant donné que, dans un grand nombre de gymnases, le programme de biologie se termine un ou deux ans avant les examens de maturité, les résultats sont divisés en deux groupes («terminé oui», «terminé non»). Les personnes testées ont, en moyenne, résolu à peine la moitié des items. Les résultats du groupe ayant terminé le programme de biologie un ou deux années plus tôt sont moins bons, tandis que ceux de l'autre groupe sont meilleurs que la moyenne. Les résultats du premier groupe en biologie sont aussi plutôt médiocres si on les compare avec ceux obtenus par ce même groupe en langue première et en mathématiques. Cela signifie que les aptitudes personnelles évaluées se situent, en moyenne, au-dessous d'un niveau moyen. Il est probable que l'oubli des connaissances joue ici un rôle qui aurait d'ailleurs pu être beaucoup plus important. Au total, aucun élève n'a obtenu le nombre maximal de points, bien que certains y soient presque parvenus. La dispersion des individus est à nouveau considérable et, dans chacun des domaines, le nombre maximal de points a été atteint. En poussant l'analyse aux valeurs agrégées sur l'ensemble des classes, on obtient également une dispersion importante des moyennes par classes. En biologie également, la meilleure classe a, en moyenne, résolu correctement trois fois plus de problèmes que la moins bonne.

L'analyse des résultats du test met en évidence une dispersion d'une amplitude étonnante dans tous les domaines et ce d'autant plus que les élèves sont sur le point d'obtenir la «maturité» qui leur donnera accès à toutes les facultés de l'université. On peut donc en déduire que la totalité des bacheliers et bachelères ne disposent pas, dans tous les domaines testés, des compétences nécessaires pour entreprendre des études universitaires dans n'importe quelle branche. Il a donc fallu s'intéresser également aux moyennes et à la dispersion des notes effectivement décernées aux examens de maturité dans les trois domaines testés.

Sauf à l'examen écrit de mathématiques, aucun élève n'a obtenu la note minimum. La dispersion relative est globalement plus restreinte que celle du test. Il apparaît en outre que la dispersion est beaucoup plus importante en mathématiques que dans les deux autres domaines. On remarque également que, pour les examens écrits de mathématiques, la note moyenne est de 4,0, ce qui correspond encore tout juste à «suffisant». En poussant l'analyse aux valeurs agrégées sur l'ensemble des classes, on obtient les valeurs extrêmes suivantes des moyennes par classe: minima: langue première, écrit 3,6, langue première, note finale 3,97, mathématiques, écrit 2,88, mathématique note finale 3,73, biologie, notation continue 3,85, sciences expérimentales, note finale

4,0; maxima: langue première, écrit 5,11, langue première, note finale 5,17, mathématiques, écrit 5,48, mathématiques, note finale 5,44, biologie, notation continue 5,28, sciences expérimentales, note finale 5,31.

Le système de notation établit clairement que les notes au-dessous de 4,0 sont insuffisantes. En analysant les résultats ci-dessus de manière plus approfondie, on constate que les pourcentages suivants de bachelères et bacheliers sont considérés comme insuffisamment qualifiés dans certains domaines selon les critères des écoles elles-mêmes (note de 3,9 ou inférieure). Examen écrit en langue première: 19,6%, note finale en langue première: 4,7%, examen écrit de mathématiques: 41,4%, note finale de mathématiques: 24%, biologie, notation continue: 5,6%, note finale en sciences expérimentales: 5,6%. Ainsi, 41% des bachelères et bacheliers se voient décerner une note insuffisante à l'examen écrit de mathématiques et 19,6% à l'examen écrit de langue première (le plus souvent une dissertation). Dans ces branches et pour la proportion de bachelères et de bacheliers indiquée ici, les compétences écrites à l'examen de maturité ont donc été jugées insuffisantes par les écoles ou par les enseignantes et enseignants concernés. En outre, on constate un saut dans la progression des chiffres lors du passage des notes insuffisantes aux notes suffisantes (note 4 ou supérieure), ce qui s'explique probablement par une tendance à arrondir vers le haut les notes situées dans cette zone. Il est donc possible que le nombre réel de qualifications insuffisantes soit encore plus élevé. L'hypothèse avancée plus haut et fondée sur les résultats d'EVAMAR II, selon laquelle les compétences des bachelères et bacheliers sont partiellement insuffisantes, est donc, du moins pour une part, le reflet de la pratique d'évaluation adoptée aussi par les examinatrices et examinateurs. Seuls les chiffres précis sont différents. Lors de la notation, il est également possible qu'une certaine adaptation au niveau de la classe joue un rôle.

Les personnes testées ont en moyenne résolu correctement un peu moins de la moitié des items du TCT dans les domaines «ensembles et formes», «compréhension de textes» et «notions élémentaires de sciences expérimentales» et un peu plus de la moitié dans le domaine «diagrammes et tableaux» uniquement. Dans l'ensemble, ces résultats se situent légèrement en dessous du test d'aptitudes pour les études de médecine (AMS). Ces valeurs inférieures peuvent s'expliquer par le fait que les bachelères et bacheliers ne se présentent de loin pas tous à l'AMS et que cette présélection volontaire induit de meilleurs résultats. La comparaison indique également que les élèves qui passaient l'examen se sont efforcés de faire de leur mieux. Les résultats des TCT constituent donc un indicateur important du sérieux avec lequel ont été effectués tous les tests. Sur l'ensemble des domaines de compétence, les aptitudes personnelles des bachelères et bacheliers correspondent, en moyenne, à un niveau légèrement supérieur à la moyenne. Ici encore, la dispersion est frappante. En poussant l'analyse aux valeurs agrégées sur l'ensemble des classes, on obtient les écarts les plus grands entre les moyennes de la meilleure et de la moins bonne classe.

Le questionnaire complémentaire portait sur l'autoévaluation des compétences transversales suivantes qui, comme les professeurs l'ont affirmé dans l'enquête, revêtent une grande importance durant les études: planification du temps de travail, préparation efficace des examens, gestion du temps, capacité de concentration, esprit d'analyse et de synthèse, capacité à argumenter, adoption d'autres perspectives. Les bachelères et bacheliers devaient indiquer sur une échelle allant de 1 (= très rarement) à respectivement 4 ou 5 (= très souvent) dans quelle mesure des affirmations relatives aux compétences transversales s'appliquaient à eux. Les chiffres montrent de fortes différences dans les résultats de l'autoévaluation des compétences transversales. Les bachelères et bacheliers indiquent en moyenne qu'ils s'astreignent à planifier leur temps «rarement à parfois» et s'estiment capables de se concentrer seulement «rarement à parfois». Ils semblent en revanche n'avoir que peu de difficulté à se préparer efficacement aux examens ou à bien gérer leur temps.

Comparaison entre strates

Pour une première comparaison, la répartition en strates a été faite selon les critères suivants: Zurich constitue une strate parce que c'est le canton le plus grand; la partie germanophone du canton de Berne, où les élèves ne doivent passer que trois ans au gymnase proprement dit pour obtenir leur maturité, forme la strate «Suisse alémanique, gym3»; «Suisse alémanique, petit» correspond aux cantons disposant de moins de 15 classes de

maturité (AI, NW, OW, GL, UR, AR, SH, VSd, ZG); «Suisse alémanique, grand» comprend les cantons disposant de plus de 15 classes de maturité (SZ, SO, TG, GR, BS, AG, SG, LU); «Suisse romande 1 (gym3)» se compose des cantons romands où la durée réglementaire du gymnase proprement dit est également de trois ans seulement (BEf, JU, NE, VD); et «Suisse romande 2» comprend les cantons romands où le gymnase se fait en quatre ans au minimum (FR, VSfr.).

En moyenne, les résultats en langue première de la strate Suisse alémanique gym3 sont significativement les moins bons. Les amplitudes de l'effet par rapport au meilleur groupe (Suisse romande 2) se situent entre très faibles et moyennes. Sur le plan des domaines de compétence, on note les différences suivantes: en «compréhension écrite générale», seuls les petits cantons alémaniques et le groupe Suisse romande 2 se démarquent des autres strates. En «compréhension écrite détaillée», on obtient à nouveau un regroupement différencié. Ici aussi, la strate Suisse romande 1 (gym3) forme un groupe à elle seule à l'extrême inférieur. Dans le domaine «fonctionnement de la langue/vocabulaire», les grands cantons alémaniques se trouvent eux aussi dans le groupe inférieur. La répartition dans le domaine «grammaire» confirme pour l'essentiel ce tableau général. Toutefois, les petits cantons alémaniques font ici encore partie du groupe des meilleurs. D'une manière générale, les gymnases dont la durée minimale est de trois ans seulement obtiennent des résultats relativement mauvais, du moins par rapport à leur région linguistique.

En mathématiques, les gymnases dont la durée minimale est de trois ans seulement obtiennent clairement de plus mauvais résultats que ceux où elle est de quatre ans, en particulier en Suisse romande. L'amplitude de l'effet entre les strates est moyenne, avec un minimum pour la Suisse romande 1 (gym 3) et un maximum pour la Suisse romande 2. On le constate encore plus nettement dans les domaines «stochastique» et «divers_m» où l'amplitude de l'effet est légèrement plus forte; en revanche, dans le domaine «géométrie», l'image est moins nette. Il semble donc que le nombre d'années passées au gymnase ait un effet encore plus important sur les résultats de mathématiques que sur ceux de langue première.

En biologie, Zurich se trouve clairement en tête avec une amplitude de l'effet importante en comparaison avec la strate obtenant la valeur la plus basse. Les gymnases de Suisse romande se regroupent vers l'extrême inférieur. Les différences entre les formations d'une durée minimale de trois ou de quatre ans dans une structure gymnasiale proprement dite ne sont pas significatives en Suisse romande, contrairement à la Suisse allemande où ces différences sont en faveur des gymnases d'une durée minimale de quatre ans.

Au test de compétences transversales, le groupe des gymnases avec une durée minimale de trois ans obtient des résultats significativement moins bons dans tous les domaines. Les amplitudes de l'effet entre les groupes situés aux extrêmes sont toutefois faibles. Il faut probablement imputer ce résultat à l'objectif même du test, qui est d'évaluer des aptitudes relativement indépendantes des connaissances acquises dans les différentes branches. La durée minimale de la formation gymnasiale au gymnase proprement dit reste certes significative, bien que dans une moindre mesure.

Les comparaisons des résultats du test dans les différentes disciplines mettent en évidence des différences moyennes à grandes entre les strates; il est dès lors pertinent de se demander si ces différences apparaissent aussi dans les notes de maturité effectivement décernées dans ces trois disciplines. Les notes de maturité en langue première présentent des différences significatives entre les strates; les amplitudes de l'effet entre les valeurs extrêmes sont moyennes. On constate cependant que le classement n'est pas identique à celui des résultats du test en langue première, phénomène auquel il n'est pas facile de trouver une explication plausible. On peut avancer l'hypothèse que la matière étudiée en langue première au gymnase est nettement plus étendue et en partie différente des connaissances mesurées à l'aide du test. Il est frappant de constater que dans l'ensemble de la Suisse romande, les notes en langue première sont inférieures à celles du reste de la Suisse. Pour la note de mathématiques, la différence entre la valeur la plus basse et la valeur la plus haute est certes significative, mais elle n'est que d'un dixième. L'amplitude de l'effet est ici très faible. En biologie, il est intéressant d'observer le classement des strates. En effet, dans le tableau des notes de maturité, il est inversé par rapport au tableau des résultats au test, avec une amplitude de l'effet presque moyenne. Les notes de maturité en

sciences ne présentent que des différences minimales entre les strates. Globalement, on ne constate pas ou très peu de différences entre les strates en ce qui concerne les notes de maturité. Cela s'explique probablement par le fait que les évaluations, à l'examen de maturité, ne sont pas et ne peuvent pas être effectuées selon une norme à l'échelle de la Suisse.

Comparaison selon l'option spécifique choisie

Les bacheliers et bacheliers de l'option spécifique Langues anciennes (surtout le latin) ont obtenu en moyenne les meilleurs résultats en langue première dans tous les domaines, ceux de l'option spécifique Arts visuels les moins bons, également dans tous les domaines. Les groupes Musique et Economie et droit figurent eux aussi, pour une partie des domaines, parmi les moins bons résultats. L'amplitude de l'effet entre les deux groupes situés aux extrêmes est moyenne. Il est frappant que les élèves ayant choisi Langues modernes n'obtiennent qu'un résultat moyen à un test de langue.

Les bacheliers et les bacheliers de l'option spécifique Physique et application des mathématiques ont clairement obtenu les meilleurs résultats en mathématiques dans tous les domaines, ceux de l'option spécifique Arts visuels les moins bons, également dans tous les domaines. Les groupes Philosophie/pédagogie/psychologie ainsi que Musique figurent eux aussi, pour une partie des domaines, parmi les moins bons résultats. Le résultat du meilleur groupe correspond à ce que l'on pouvait attendre. L'amplitude de l'effet entre ce dernier et le moins bon groupe est toutefois exceptionnellement forte. Les élèves du groupe Langues anciennes se classent au deuxième rang, devant ceux du groupe Biologie et chimie et ceux du groupe Economie et droit. L'amplitude de l'effet du groupe en deuxième position Langues anciennes par rapport au moins bon reste considérable.

Comme on pouvait s'y attendre, les bacheliers et bacheliers ayant choisi l'option spécifique Biologie et chimie ont obtenu les meilleurs résultats en biologie dans tous les domaines (voir tableau 15). A ce test aussi, les élèves du groupe Langues anciennes se classent au deuxième rang dans tous les domaines et devancent même le groupe Physique et application des mathématiques. Les moins bons résultats globaux sont ceux du groupe Musique tandis que les groupes Langues modernes, Philosophie/pédagogie/psychologie et Economie et droit sont au bas du classement pour certains domaines. Les amplitudes de l'effet entre le moins bon groupe et le meilleur sont très fortes. Dans la comparaison entre le plus mauvais et le deuxième groupe, Langues anciennes, elles restent à un niveau modéré.

Le classement des groupes au test de compétences transversales est le même dans tous les domaines. Les différences entre le meilleur groupe (Physique et application de mathématiques) et le moins bon (Philosophie/pédagogie/psychologie) sont considérables (très forte amplitude de l'effet). Les élèves de l'option Arts visuels se situent dans le groupe de signification ayant les plus mauvais résultats.

La comparaison des résultats entre les groupes selon l'option spécifique choisie ayant mis en évidence des différences parfois importantes, il paraissait intéressant d'examiner si de telles différences apparaissent aussi en ce qui concerne les notes de maturité dans les trois disciplines. L'amplitude de l'effet entre les groupes situés aux extrêmes reste moyenne dans la comparaison concernant la note de maturité en langue première. Il faut relever que les élèves de l'option spécifique Musique sont mieux placés que dans le classement des résultats au test, tandis que ceux de l'option spécifique Biologie et chimie reculent. En ce qui concerne la note de maturité en mathématiques, on observe que l'amplitude de l'effet mesurée entre les groupes situés aux extrêmes diminue presque de moitié, bien que restant assez élevée. Pour cette discipline, le reste du classement correspond bien à celui des résultats au test. Il semble donc que les personnes chargées d'évaluer les élèves en mathématiques réussissent mieux à appliquer des critères objectifs. En biologie (notation continue), la comparaison des notes de maturité montre que l'amplitude de l'effet entre les groupes situés aux extrêmes diminue à une valeur moyenne. Le classement correspond étonnamment bien, à quelques écarts près, à celui du test de biologie. Les plus grands

écarts constatés concernent les élèves des options Biologie et chimie et Arts visuels. Les notes des premiers sont légèrement moins bonnes (ils sont d'ailleurs classés derrière les élèves de Langues anciennes) tandis que celles des seconds sont meilleures.

En complément aux comparaisons de moyennes ci-dessus, les pourcentages de résultats insuffisants ont été calculés. En mathématiques, la moitié des élèves obtiennent des résultats insuffisants à l'examen de maturité; cette proportion est même un peu plus élevée pour ceux des nouvelles options spécifiques Philosophie/pédagogie/psychologie, Arts visuels et Musique créées après le RRM 95. En outre, il apparaît que la notation continue a l'effet de compenser le relativement grand nombre d'évaluations insuffisantes aux examens écrits.

Comparaison entre les sexes

Comme dans la plupart des recherches, nous avons procédé à une comparaison des résultats obtenus par chacun des sexes aux tests et aux examens de maturité. Sur l'ensemble des comparaisons, il s'avère que la majorité des bachelières obtiennent de meilleurs résultats que les bacheliers en langue première (différence minime sur les notes de test), mais de moins bons en mathématiques, en biologie et au test d'aptitudes transversales. Il faut relever le contraste entre le résultat significativement moins bon des bachelières à tous les tests de biologie et la notation continue en biologie significativement meilleure de ces dernières. On constate également que les bachelières ont compensé leurs résultats plus mauvais en physique par de meilleures notes en biologie (les différences dans la notation continue en chimie sont négligeables).

Comparaisons selon les taux de maturité gymnasiale

En Suisse, les taux de maturités gymnasiales diffèrent fortement d'un canton à l'autre. D'après les statistiques fédérales, 29,5% des jeunes avaient obtenu une maturité au Tessin en 2007, tandis qu'ils n'étaient que 11,2% dans le canton d'Obwald. Ces différences ne reflètent pas nécessairement une inégalité des chances. Il est toutefois intéressant d'examiner si les mêmes exigences sont posées aux examens de maturité. EVAMAR II permet de comparer certaines des compétences des bachelières et bacheliers provenant de cantons dont les taux de maturités gymnasiales sont différents. L'échantillon n'étant pas représentatif pour chaque canton, les taux de maturités n'ont pu être répartis qu'en catégories assez grossières. Des superpositions avec d'autres effets restent donc possibles, et la question de l'égalité des exigences lors de la sélection à l'entrée au gymnase reste à traiter. Lors de l'évaluation de l'égalité des chances, il serait nécessaire de prendre en compte cet élément car un «rattrapage» des déficits de compétence durant les années de gymnase est tout à fait possible et aurait pour effet d'annuler à la fin de cette période les différences qui existaient au début, sans pour autant supprimer l'inégalité des chances. En raison de la superposition possible des effets, nous présentons ci-après une comparaison entre un groupe composé du tiers supérieur et un groupe composé du tiers inférieur. Les bachelières et bacheliers du groupe dont le taux de maturités est le plus bas ont obtenu les meilleurs résultats à presque tous les tests et épreuves. Les différences les plus grandes s'observent dans les résultats en biologie (amplitude de l'effet moyenne, au minimum). Pour la plupart des notes de maturité, les différences, bien que moindres, vont dans le même sens que celles observées dans les résultats du test. Les notes de langue première font exception car les différences qu'elles présentent dépassent même celles des résultats du test. On peut donc en conclure que les taux de maturités élevés sont en majeure partie liés à des résultats moins bons dans les parties de la formation gymnasiale évaluée par EVAMAR II et que, en conséquence, il faut se poser la question de l'égalité des chances.

Comparaison selon l'élément structurel: gymnase de courte ou longue durée

En Suisse alémanique, une partie des gymnases commencent directement après le degré primaire, c'est-à-dire que le degré secondaire I est intégré aux études gymnasiales, qui sont ainsi plus longues. Il était donc

particulièrement intéressant d'observer si les résultats présentaient des différences entre ces deux systèmes. La comparaison montre que, dans presque tous les domaines testés ainsi qu'aux notes de maturité, les gymnases longue durée obtiennent des résultats significativement meilleurs. Cependant, l'amplitude de l'effet reste faible dans la langue première (sauf dans la partie «grammaire») ainsi qu'en mathématiques. En ce qui concerne les tests de biologie (mais pas les notes), l'amplitude de l'effet se situe déjà dans un domaine moyen. Pour expliquer ces observations, on peut supposer que les disciplines enseignées de manière systématique comme la biologie (mais aussi la chimie, la physique, l'histoire, etc.) bénéficient du fait qu'elles sont très structurées dès le degré secondaire I, que leur enseignement est assuré par une personne ayant une formation universitaire spécialisée et qu'il est planifié sur toute la durée du gymnase si bien que le passage de l'école secondaire au gymnase de courte durée se fait sans perte d'efficacité dans l'apprentissage malgré la grande hétérogénéité des conditions préalables requises dans ces disciplines.

5 Résultats III: Analyse des examens de maturité

Le sous-projet D1 d'EVAMAR II était consacré à l'analyse qualitative des examens écrits de maturité des bacheliers et bacheliers du même échantillon que le module C (y compris le Tessin). Il s'agissait d'évaluer les exigences posées aux élèves sur le plan cognitif et sur celui du contenu, et d'examiner dans quelle mesure elles correspondent, d'une part, aux différentes composantes de l'aptitude aux études supérieures identifiées dans le cadre des modules A et B et, d'autre part, aux tests menés dans le module C. Cette analyse a été effectuée en premier lieu pour les disciplines ayant fait l'objet des tests: langue première, mathématiques et biologie; on pourrait envisager de l'étendre à d'autres disciplines lors d'un nouveau projet.

En Suisse, les examens écrits de maturité en langue première se distinguent par leur extraordinaire diversité: il semblerait que chaque canton, chaque école, chaque enseignante ou enseignant même, ait sa propre idée de la manière de composer cet examen et des conditions permettant de vérifier si les bacheliers et bacheliers possèdent la «maturité nécessaire aux études universitaires». Le corps enseignant suisse est généralement d'accord sur deux points: la majeure partie des textes sont dans leur langue originale (il s'agit rarement de traductions) et les auteurs féminins sont rarement choisis. Les enseignantes et enseignants ont plutôt tendance à choisir un ou deux textes fondamentaux du 20^e ou du 21^e siècle sur un sujet aussi récent que possible ou un thème intemporel. Ils choisissent en l'occurrence dans 40 à 60% des cas des textes littéraires et seulement dans 10 à 20% des cas des textes documentaires. Dans les autres cas, il s'agit de textes dits «philosophiques» (allemand 12%, français 14%, italien 30%). Le niveau d'exigence de la plupart des examens écrits de maturité se situe dans la partie supérieure (niveaux 4-6 selon Anderson et al.) des exigences cognitives (D 77%, F 99,1%, I 77%). En ce qui concerne les principales différences régionales, on constate une nette prédominance du niveau d'exigence «création – critique» en Suisse romande (deux tiers), alors qu'il ne vient qu'en troisième place au Tessin (un cinquième). En Suisse alémanique, la répartition est plus équilibrée: la réflexion critique forme un tiers des exigences, l'analyse un quart et la créativité un autre quart. Globalement, la dissertation de quatre heures constitue un exercice du plus haut niveau cognitif, celui de la création. Sur le plan du contenu, les thèmes soumis aux élèves se réfèrent au monde actuel, plus particulièrement à celui des jeunes, et mettent l'accent sur la diversité des opinions et les débats de société ou les controverses sur des questions existentielles. Si l'on se réfère à l'enquête auprès des professeurs, il apparaît que les multiples contenus d'examens (auteurs, thèmes de société, problèmes éthiques, etc.) ne correspondent que très peu aux attentes des professeurs interrogés, toutes branches confondues. De surcroît, la compréhension écrite est évaluée essentiellement à partir de textes littéraires et très peu à partir de textes scientifiques.

Dans les examens de maturité de biologie, 40,8% des copies analysées contenaient des *higher order questions* (degrés principaux 3-6 chez Anderson et al.). Cette valeur permet de les comparer avec les assessments

américains bien connus, tels que le *Medical College admission Test* (MCAT) ou l'*Advanced Placement (AP) Biology Course*. Et cela ne va pas de soi pour deux raisons:

- a) Dans les années 1980 seulement, soit nettement plus tard que dans le contexte américain, le monde germanophone a commencé à s'intéresser à la pensée productive. L'expérience montre qu'il faut souvent plusieurs dizaines d'années pour que les conclusions des recherches universitaires, même bien étayées empiriquement, fassent leur apparition dans la formation des enseignantes et enseignants. Les présents résultats semblent indiquer que ce transfert a largement été opéré dans la formation didactique des enseignantes et enseignants du secondaire II en Suisse, et qu'il porte ses fruits dans bon nombre de salles de classe des gymnases.
- b) Créer et corriger un bon exercice de transfert est une tâche longue et compliquée. Manifestement, beaucoup d'enseignantes et enseignants sont persuadés que le jeu en vaut la chandelle.

Le pourcentage élevé de *higher order questions* contredit également le cliché selon lequel, en biologie, il suffit d'apprendre par cœur sans réfléchir pour obtenir une bonne note. Les résultats de l'analyse des examens de biologie indiquent toutefois aussi que les domaines «écologie et évolution» pourraient avoir été insuffisamment traités.

Les examens de mathématiques présentent eux aussi une grande diversité. En règle générale, les épreuves de maturité sont élaborées avec soin et beaucoup d'imagination. Il existe cependant aussi un nombre considérable d'épreuves «standard» avec un problème par matière (analyse de courbes, calcul vectoriel, calcul de probabilité, calcul différentiel) et un exercice subdivisé en plusieurs items pour couvrir les autres domaines (suites et séries, algèbre élémentaire, etc.) En ce qui concerne le niveau d'exigence cognitif, le pourcentage de mémorisation est particulièrement fort dans ces examens. Il existe de grandes différences sur le plan des exigences cognitives (potentielles) tant entre les examens de mathématiques comme discipline fondamentale et les examens de l'option spécifique qu'entre les examens de type discipline fondamentale pratiqués par les diverses écoles. En ce qui concerne le potentiel d'exigences cognitives, la majeure partie des épreuves (partielles) de maturité appartiennent à la catégorie «exécuter: appliquer une procédure à un problème déjà connu». Le problème d'extrema apparaissant dans presque chaque examen appartient généralement, à notre avis, à la catégorie «résoudre un nouveau problème en le ramenant à un problème connu». Certains problèmes d'extrema plus difficiles nécessitent un important travail de modélisation, ce qui correspond au niveau «modéliser: utiliser des modèles mathématiques connus» (au cas où le modèle est déjà connu) ou au niveau «modéliser: utiliser de nouveaux modèles mathématiques». Ces types de problèmes n'apparaissent que dans les examens destinés aux élèves de l'option spécifique Physique et applications des mathématiques.

6 Résultats IV: Analyse des travaux de maturité

Ce module a été entièrement conçu et réalisé par la Haute école pédagogique de la Fachhochschule Nordwestschweiz.

L'analyse de la qualité des travaux de maturité en Suisse alémanique, en Suisse romande et au Tessin se fondait sur trois éléments: l'analyse de 437 travaux de maturité, notés chacun par deux experts indépendants, qui formait le volet principal de l'enquête; un questionnaire invitant un échantillon de bacheliers et bachelères (incluant les auteurs de ces travaux) à s'exprimer par écrit sur le contexte, le déroulement et les résultats de leur projet; et une reconstitution des conditions générales institutionnelles, réalisée aux niveaux national et cantonal et à celui des 30 gymnases ayant pris part à l'enquête.

L'évaluation des travaux de maturité conclut dans l'ensemble à un bilan positif. Il s'avère en effet que la grande majorité des bacheliers et bachelères sont familiarisés avec les règles fondamentales du travail scientifique.

L'analyse confirme l'hypothèse que la qualité du contenu des travaux est en corrélation avec leur qualité

formelle et leur niveau de langue: les élèves bien notés pour les premiers critères le sont généralement aussi pour les seconds, tandis que les travaux dont le contenu laisse à désirer présentent également davantage de lacunes que la moyenne sur les plans de la forme et de la langue. Le fait de structurer le choix du thème traité en posant des cadres thématiques a, d'autre part, un effet positif sur la qualité des travaux de maturité: dans les écoles ne pratiquant pas cette délimitation, les travaux sont en moyenne un peu moins bons. Il est vrai que plus le thème est large, plus les tâches d'encadrement sont complexes.

On ne constate pas de grandes différences de qualité dans ces travaux entre les sciences expérimentales, sociales et humaines ni, du reste, entre les sexes. Il en va autrement de la comparaison entre les régions: même si le bilan général est positif pour toutes les parties du pays, on ne peut manquer de voir certaines différences considérables entre la Suisse alémanique et la Suisse latine, qui n'ont pu obtenir d'explication statistique concluante dans le cadre de ce projet. On peut supposer que les meilleurs résultats de la Suisse romande et du Tessin sont dus à des consignes plus étroites quant au choix thématique et au fait que le travail s'inscrivait dans des structures de type séminaire. Mais on ne peut exclure que les résultats de la notation aient été influencés par des cultures d'évaluation différentes d'une région linguistique à l'autre.

S'agissant des conditions générales, il faut admettre que les bacheliers et bachelières utilisent peu les guides et autres documents officiels de leur établissement; c'est plutôt sur la base des contacts interpersonnels avec les enseignants qui les encadrent qu'ils se font une idée claire et précise de ce que l'on attend d'eux. A noter que l'intensité avec laquelle ils mettent à profit le soutien apporté dans la gestion des phases critiques est plus bien déterminante que la nature et la fréquence de ces contacts. Cela va de la formulation du thème à la construction formelle du rapport en passant par la réponse aux questions méthodologiques. A l'image de la fréquence des contacts, la durée de réalisation des projets n'a que peu d'impact sur la qualité des travaux de maturité.

Ce qui frappe dans l'analyse, c'est que la notation des experts est généralement plus sévère que celle du corps enseignant. Il faut toutefois tenir compte du fait que l'objet de l'évaluation pédagogique diffère sous plusieurs aspects de celui sur lequel s'est penchée la présente enquête. Les enseignantes et enseignants ne jugent en règle générale pas uniquement le produit écrit, mais également la présentation orale. La présente analyse qualitative a, de plus, dû se limiter à une dimension sommative, puisqu'elle n'a pu se faire tout au long du processus, tandis que les enseignantes et enseignants ont aussi sous les yeux les progrès de leurs élèves tout au long du déroulement du projet. De surcroît, il convient de souligner que l'évaluation des travaux de maturité réalisée dans le cadre de cette enquête se borne dans une large mesure à la question académique de l'acquisition de l'aptitude aux études supérieures, alors que les enseignantes et enseignants, par leur mandat éducatif, peuvent faire intervenir une multitude d'aspects différents dans leur appréciation.

Les élèves ont tendance à attribuer au travail de maturité une valeur à la fois intrinsèque et extrinsèque; les cas sont relativement peu nombreux où l'exercice est rétroactivement jugé inutile. La qualité du contenu a une corrélation positive avec le regard porté par les élèves sur l'utilité de ces travaux pour la suite de leur parcours de formation académique ou professionnelle comme pour leur développement personnel.

Du point de vue des institutions tertiaires, l'image que l'on se fait à première vue du travail de maturité est que les bacheliers et bachelières devraient être en mesure de fournir un travail satisfaisant aux exigences élémentaires d'un cursus de haute école. Tant la structuration des textes que l'utilisation de citations et la présence de bibliographies dans les travaux de maturité témoignent que les règles académiques courantes sont déjà exercées au gymnase. Il en va de même pour le compte-rendu de faits théoriques et empiriques élémentaires et, à l'état d'ébauche, pour l'obtention d'informations procédant par recherche ou par expérimentation, même si la part de description et de reproduction est nettement plus importante dans les projets scolaires que l'approche analytique.

Il faut toutefois souligner à ce sujet qu'une analyse transversale des travaux de maturité ne permet pas de faire la moindre prévision quant à la durabilité des processus d'apprentissage. La haute qualité des résultats est favorisée par la synergie de facteurs contextuels tels que la portée des mesures institutionnelles préventives, l'intensité de l'encadrement pédagogique et le soutien apporté par l'environnement social. Si ces facteurs ne sont

plus réunis dans le cadre de la vie estudiantine, on ne peut exclure que bon nombre des premiers travaux des nouveaux étudiants se révèlent d'un niveau inférieur à celui des travaux de maturité.

7 Les limites de l'étude EVAMAR II

Il convient de rappeler encore une fois expressément les limites de l'étude EVAMAR II déjà évoquées au point 2. Le projet permet de commenter seulement certains aspects choisis des objectifs de la formation gymnasiale. Voici les éléments relevant d'une aptitude générale aux études supérieures qu'EVAMAR II n'a pas étudiés et dont il n'a pas analysé la présence chez les bacheliers et bacheliers: premièrement les connaissances et compétences qui ne sont pas directement exigées à l'université, bien qu'elles fassent partie d'une vaste culture générale, permettant notamment d'assumer ses responsabilités au sein de la société (maturité sociale), et qui doivent donc s'acquérir dans le cadre de la formation gymnasiale; deuxièmement les connaissances et compétences qui font nécessairement partie de l'aptitude aux études supérieures mais qui sont issues d'autres disciplines que la langue première, les mathématiques et la biologie; et troisièmement les éléments de compétences transversales que le cadre défini pour le projet n'aurait permis de saisir qu'imparfaitement, malgré leur importance pour la réussite des études (par ex. la capacité d'apprendre de manière autonome, de s'organiser de manière générale, de faire des recherches, d'assimiler rapidement une riche bibliographie spécialisée, etc.). La portée de ces facteurs ne pourrait être établie qu'à l'aide d'autres enquêtes, proposées dans le concept général de l'étude (études longitudinales auprès des étudiants) et il faudrait pour la mesurer des procédures d'évaluation lourdes à réaliser. Rappelons toutefois que ces facteurs ont déjà été étudiés et validés approximativement dans le cadre du projet EVAMAR I par le biais de questionnaires, autrement dit par autoévaluation. Le questionnaire joint à EVAMAR II comprenait en outre des questions à ce sujet complémentaires à EVAMAR I.

VII Conclusions et recommandations du chef de projet

Le présent chapitre présente des conclusions et des recommandations qui paraissent importantes au chef de projet. On notera qu'elles comportent des appréciations à caractère normatif.

1 Des résultats généraux satisfaisants

Les bachelières et les bacheliers suisses ont obtenu, dans les domaines étudiés par le projet EVAMAR II, des résultats normaux pour des tests standardisés mis au point pour une population déterminée. Il est dans la norme que les items aient été résolus correctement pour moitié environ. Les aptitudes des bachelières et des bacheliers se situent en moyenne à un niveau permettant de résoudre correctement des problèmes de difficulté moyenne par rapport aux critères universitaires. Dans ce sens, on peut qualifier de satisfaisant le niveau de formation des bachelières et des bacheliers suisses.

Les possibilités et les limites de la mesure des performances, et plus particulièrement du modèle théorique de test choisi, excluent la construction d'un test dans lequel il faudrait que la *totalité* des items soient résolus correctement pour faire état de la compétence mesurée, en l'occurrence certains aspects de l'aptitude aux études supérieures. Indépendamment du fait qu'il serait quasi impossible de procéder à l'évaluation d'un tel test en employant les procédures connues dans la théorie des tests scientifiques, les items proposés seraient jugés trop faciles et ne seraient pas pris au sérieux par les observateurs extérieurs. Mais même des résultats «normaux» ne permettent pas de déterminer avec précision à partir de quel nombre de points l'aptitude aux études supérieures est donnée dans le domaine testé. Les seuils fixés pour réussir des examens et être admis à l'université, par exemple à travers les critères de réussite des examens de maturité ou les examens d'admission, n'offrent pas non plus cette possibilité. En effet, il n'existe pas de méthode capable de fournir une mesure précise et déterministe de caractéristiques psychologiques. De plus, l'aptitude aux études supérieures n'est pas une qualité absolue et stable que l'on peut cerner par une méthode dichotomique. Par conséquent, il est impossible de faire un calcul déterministe indiquant où se situe le seuil de performance dans les tests d'EVAMAR II au-dessous duquel il devient impossible de réussir des études supérieures.

Pour faire des prévisions probabilistes sur la réussite universitaire de personnes en utilisant des instruments de mesure psychométriques, il faut au préalable avoir établi la validité prévisionnelle de ces instruments de mesure par comparaison avec la réussite effective des personnes étudiées. Or, de telles études s'étalent nécessairement sur plusieurs années, les points de mesure étant très éloignés dans le temps. Mais même dans ces conditions, il n'est pas possible de faire des prévisions à caractère déterministe; les études réalisées à ce jour ont établi des corrélations ne dépassant pas un coefficient de 0,5 (lire les descriptions correspondantes aux chapitres I et IV). Les délais étaient trop courts pour que l'on puisse entreprendre une validation prévisionnelle des procédés de mesure mis en œuvre dans EVAMAR II. Mais la solidité de leur fondement théorique et le fait que les tests ont été développés avec systématisme pour l'interface entre gymnase et université permettent de fournir des interprétations fondées des résultats du test allant plus loin que le simple constat de «normalité» exposé plus haut. Si l'on voulait des interprétations plus poussées et, par exemple, pouvoir déterminer quel nombre de points il faut avoir obtenu dans les différents tests pour avoir telle probabilité de réussir des études supérieures dans telle filière, il faudrait procéder à des études empiriques supplémentaires exhaustives et fastidieuses.

La grande dispersion observée dans les résultats au test signifie également que tous les gymnases ne libèrent pas leurs bachelières et leurs bacheliers avec des compétences que l'on peut qualifier au minimum de suffisantes dans toute leur étendue. Cela confirme ce qui n'était encore qu'une supposition au vu des notes insuffisantes effectivement décernées dans les gymnases aux examens de maturité, surtout en mathématiques et dans l'épreuve écrite de la langue première. De même, les taux d'échec plus ou moins élevés aux examens universitaires s'expliquent en partie par le fait qu'une partie des étudiantes et étudiants recalés ne possédaient

pas d'aptitude générale aux études supérieures. Mais cela signifie aussi que, malgré le niveau de formation satisfaisant des bachelières et des bacheliers suisses, il existe au fond un potentiel d'amélioration. Cette conclusion n'a rien d'extraordinaire; elle est valable à tous les degrés scolaires dont on a mesuré le niveau des élèves au moyen de tests standardisés. On ajoutera qu'aucun degré scolaire sélectif ne parvient à libérer ses élèves de manière à assurer la réussite de tous au degré suivant.

Des mesures s'imposeraient si bien moins de la moitié des items avaient été résolus correctement et si, en moyenne, le niveau des aptitudes personnelles était clairement inférieur à une note moyenne. Ce n'est pas le cas des bachelières et des bacheliers suisses.

Quelle est la proportion de connaissances que l'on retient ou que l'on oublie à la sortie d'un cycle de formation? Cette question n'a pas été bien étudiée. Les bachelières et les bacheliers qui avaient terminé le programme de biologie plus tôt n'avaient pas tout oublié, mais ils ont obtenu des résultats significativement moins bons dans le test de biologie. On peut supposer que leurs connaissances et compétences au moment où le programme de biologie s'est achevé étaient plus importantes qu'au moment du test, qui a eu lieu juste avant les épreuves de la maturité. Si donc les universités déplorent parfois des lacunes dans les connaissances des nouveaux étudiants et étudiantes, la cause de ces lacunes ne réside pas obligatoirement dans le (trop) faible nombre d'heures accordées à une discipline au gymnase, mais elle est peut-être due au fait que l'enseignement de cette discipline s'arrête trop tôt. Il est donc souhaitable que l'enseignement de toutes les disciplines se poursuive aussi tard que possible, même si le maintien d'une grande variété de matières jusqu'à la maturité comporte d'autres inconvénients.

2 Assurer l'aptitude générale aux études supérieures

Si l'on comprend l'aptitude générale aux études supérieures comme supposant que chaque certificat de maturité doit attester la possession des compétences suffisantes pour entreprendre des études supérieures dans n'importe quelle branche, on peut présumer qu'une partie considérable des bachelières et des bacheliers en sont dépourvus. Le projet EVAMAR II a fait ce constat seulement dans les domaines testés, mais rien ne permet de penser que l'intégration d'autres disciplines dans le test aurait eu un impact sur la proportion importante de performances insuffisantes. On ne peut toutefois pas dire avec précision quel pourcentage représente cette proportion (voir les considérations exposées à la section 1).

Ce constat n'a rien de nouveau et, comme évoqué plus haut, les lacunes observées dans le test se reflètent dans les notes de maturité en langue première et plus encore en mathématiques. Lorsque les bachelières et les bacheliers obtiennent une note insuffisante dans les épreuves écrites de la maturité dans un cinquième des cas (langue première) voire dans plus de deux cinquièmes des cas (mathématiques), il est évident qu'ils présentent des lacunes considérables dans ces disciplines. C'est le système de compensation des notes introduit par le RRM qui permet malgré cela à l'ensemble des bachelières et des bacheliers d'obtenir le certificat de maturité qui atteste l'aptitude générale aux études supérieures. Cela explique pourquoi les professeurs d'université interrogés dans le cadre du sous-projet A4 se plaignent de lacunes surtout en langue première et en mathématiques et en partie pourquoi les taux d'échec aux premiers examens sont si élevés dans certaines branches d'études. Reste à déterminer si les éléments de connaissances et de compétences à l'origine des résultats insuffisants aux examens de maturité sont indispensables pour avoir une aptitude générale aux études supérieures. A la limite, des bachelières et des bacheliers ayant des notes insuffisantes pourraient malgré tout avoir un niveau suffisant dans des domaines pertinents pour l'aptitude générale aux études supérieures ou, inversement, les bachelières et les bacheliers ayant des notes suffisantes pourraient ne pas avoir un niveau suffisant dans des domaines de la discipline pertinents pour l'aptitude générale aux études. Les corrélations au mieux moyennes observées entre les résultats du test et les notes de maturité montrent également que les secondes ne mesurent que partiellement la même chose que les tests employés dans le cadre d'EVAMAR II (voir également les comparaisons au chapitre V). Néanmoins, les élèves ont obtenu des résultats en partie mauvais dans les tests portant sur l'ensemble des connaissances et des compétences pertinentes pour au moins une branche d'études.

Si l'on tient compte en outre des notes de maturité, on a là des raisons fondées de supposer que tous les bacheliers et bachelières ne possèdent pas un niveau suffisant dans toute l'étendue des compétences nécessaires pour entreprendre les études supérieures de leur choix sans avoir à combler des lacunes. Et pourtant, ils ont été qualifiés de généralement aptes aux études supérieures puisqu'on leur a décerné le certificat de maturité.

La question de fond est donc la suivante: faut-il se donner pour but que toutes les bachelières et tous les bacheliers possèdent à l'avenir une «aptitude sans lacune aux études supérieures»? Bien qu'il n'existe pas d'instruments de mesure adaptés à coup sûr à ce but, on peut envisager de l'atteindre grâce à différentes mesures, qu'il convient de mentionner et de discuter ici.

- Variante 1: On pourrait réviser les critères de réussite des examens de maturité de façon à ne plus autoriser la compensation des notes insuffisantes.

Mais cela aurait vraisemblablement deux effets indésirables. Dans la mesure où l'on ne peut pas améliorer à volonté les performances cognitives des élèves et où les gymnases sélectionnent vraisemblablement déjà les meilleurs sur ce plan, la suppression de la compensation des notes avec le maintien des critères d'évaluation actuels ferait chuter le taux de maturité. On ne peut pas s'attendre à ce que les 25% de notes insuffisantes en mathématiques observées actuellement se transforment en notes suffisantes. Pour éviter une baisse aussi dramatique, il faudrait adapter la notation en assouplissant les critères d'évaluation. Deuxièmement, il n'est pas sûr que la qualification exprimée par les notes attribuées dans les différentes disciplines soit équivalente à la qualification des compétences qu'il faut posséder dans ces disciplines pour être apte à entreprendre des études supérieures (voir les réflexions à ce propos dans le 2^e paragraphe ci-dessus). Rappelons ici un point important qu'il importe de garder à l'esprit: l'acquisition de l'aptitude aux études supérieures n'est pas le seul but de la formation gymnasiale.

- Variante 2: On pourrait élargir l'approche méthodologique d'EVAMAR II pour déterminer quelles sont les compétences dans les différentes disciplines gymnasiales qui sont véritablement importantes pour l'aptitude générale aux études supérieures puis déclarer que les résultats insuffisants dans ces compétences ne sont pas compensables. Il faudrait alors que ces compétences soient évaluées à part dans les examens de maturité.

Indépendamment du travail énorme que cette étude demanderait (de l'analyse des exigences jusqu'à l'élaboration de tests), des difficultés méthodologiques qu'elle présenterait et des éventuels signaux indésirables qu'elle adresserait aux apprenantes et aux apprenants («dévalorisation» des domaines jugés moins pertinents pour l'aptitude aux études supérieures), un tel critère de réussite aboutirait vraisemblablement lui aussi à une baisse du taux de maturité politiquement indésirable.

L'objectif d'atteindre une «aptitude sans lacune aux études supérieures» pour l'ensemble des bachelières et des bacheliers avec un taux de maturité inchangé est très probablement une fiction. Il ne constitue pas non plus, à notre avis, une nécessité absolue car bon nombre de bachelières et de bacheliers sont probablement tout à fait en mesure d'acquérir ultérieurement les connaissances et compétences qui leur manquent de manière limitée dans certains domaines au moment de l'obtention de la maturité, en fonction de leurs besoins et de l'orientation choisie. La seule condition à cela est qu'ils possèdent la capacité d'apprendre de manière autonome.

- Variante 3: On pourrait envisager d'abandonner l'exigence de l'aptitude générale aux études de même que le droit général des titulaires d'une maturité suisse à l'admission aux universités. Cela reviendrait à se limiter à une maturité «spécialisée», différentes variantes concrètes étant envisageables pour son acquisition au gymnase comme pour les droits d'admission qu'elle donne effectivement. Au niveau du gymnase, on pourrait par exemple envisager un système à deux étages, dans lequel la spécialisation interviendrait seulement en deuxième partie de scolarité, pendant la dernière ou les deux dernières années gymnasiales. Le certificat de maturité ne donnerait alors le droit d'être admis que dans une faculté déterminée. On pourrait aller plus loin en prescrivant une spécialisation formelle dans le domaine d'études supérieures visé, ce qui est le cas dans

certaines systèmes anglo-saxons. Il serait alors clair que l'on ne peut pas entreprendre des études dans un domaine sans avoir un niveau suffisant dans celui-ci.

Dans ce type de solution, la définition des exigences propres à une faculté ne serait pas sans poser de problème. Les enquêtes réalisées dans le cadre du module A ont montré qu'il n'existe pas de lien structurel linéaire entre les disciplines de maturité et les branches d'études supérieures, mais que chaque branche d'études requiert un socle de connaissances et de compétences gymnasiales dans un éventail plus ou moins large de disciplines. Dans ces conditions, il paraîtrait naturel de confier à l'université la compétence de faire passer les examens et d'accorder les droits d'admission. Mais une maturité spécialisée amputerait la formation gymnasiale de ses autres buts, notamment la «préparation à assumer des responsabilités au sein de la société actuelle» (voir section suivante). Il ne faut donc pas renoncer à l'ambition de décerner un certificat de maturité qui atteste la possession de la maturité nécessaire pour entreprendre des études universitaires et donne le droit d'accéder sans examen à toutes les branches d'études supérieures.

Les restrictions qui s'appliquent actuellement aux domaines de compétence dans lesquels un niveau insuffisant est autorisé constituent un argument important en faveur du maintien du droit général d'admission aux universités malgré des lacunes dans certains domaines. Les bacheliers et les bacheliers ayant des résultats insuffisants doivent faire état de résultats deux fois meilleurs dans d'autres disciplines et ils ne peuvent pas avoir plus de trois notes inférieures à 4. Selon les résultats d'EVAMAR I, ce changement par rapport à la réglementation antérieure à la réforme a entraîné un quasi-doublement du taux d'échec, tout au moins en Suisse alémanique (Ramseier et al., 2005, p. 139).

Il faut en outre tenir compte des effets dans la pratique des lacunes constatées: le projet EVAMAR I et d'autres études ont mis en évidence que la canalisation du choix des études en fonction des possibilités individuelles (et des intérêts associés) se fait automatiquement si bien que les lacunes dans certaines disciplines n'ont qu'un impact limité sur la réussite concrète des études supérieures.

- Variante 4: Il faut se demander s'il ne serait pas judicieux, dans le cadre d'une révision du RRM, de définir des critères de réussite sans possibilité de compenser les résultats insuffisants, qui seraient applicables au minimum à une sélection de compétences de base et pour l'ensemble des bacheliers et des bacheliers. Compte tenu des résultats d'EVAMAR II, cela concernerait les compétences en langue première et en mathématiques. Les compétences de base ne couvrent pas toute l'étendue du curriculum de chaque discipline gymnasiale; il s'agit des compétences pertinentes pour une multiplicité de domaines d'études. Elles ne doivent donc pas être mises sur le même plan que les compétences en langue première et en mathématiques mesurées lors des examens de maturité, dont elles ne constituent qu'une portion. La possibilité de compenser des notes insuffisantes serait maintenue pour les autres éléments du programme d'études, sur lesquels l'examen de maturité porte également car ils font partie des objectifs du gymnase et sont tout aussi importants. Il ne s'agit donc pas d'exiger de manière générale une note de maturité minimale de 4 en mathématiques et en langue première. Comme dans la variante 2, les épreuves portant sur ces domaines lors des examens de maturité devraient être évaluées à part. Toutefois, les analyses d'EVAMAR II ne permettent pas de déterminer avec suffisamment de précision en quoi consistent les compétences de base. Il faudrait donc préciser leur définition conformément aux règles de l'élaboration des modèles de compétence. A cet effet, il faudrait compléter la procédure utilisée pour HarMoS d'une part en intégrant la démarche d'analyse pratiquée dans EVAMAR II pour déterminer les connaissances et les compétences requises en langue première et en mathématiques et, d'autre part, en organisant un dialogue entre les acteurs déterminants à l'interface entre gymnase et université afin d'aboutir à un consensus, qui aurait un caractère obligatoire, sur les compétences de base souhaitées et nécessaires.

Le volume de travail redouté pour la variante 2 subsisterait dans ce cas, mais il ne devrait plus être fourni pour toute l'étendue de l'ensemble des disciplines de maturité. L'impossibilité de compenser des résultats insuffisants dans un petit nombre de domaines n'englobant pas des disciplines entières devrait en outre avoir un impact supportable sur le taux de maturité s'il devait baisser.

Le chef de projet est partisan de la variante 4. Elle est de nature à conforter la maturité dans son rôle d'attestation de l'aptitude générale aux études. On peut également penser que des critères de réussite sans possibilités de compensation produiraient un meilleur effet de guidage dans la mesure où les bacheliers et les bacheliers feraient moins vite des calculs susceptibles de les conduire à une résignation anticipée dans ces domaines. Comme l'anglais en tant que langue de la science fait partie des compétences de base de l'aptitude aux études supérieures – ce qu'estimaient d'ailleurs les professeurs interrogés dans le cadre du sous-projet A4 –, il faudrait imposer de la même manière un niveau de compétence standard en anglais (qui reste à déterminer).

3 Maintenir un large éventail de disciplines et le droit général à l'admission aux universités

Le chef de projet recommande, à la section 2, de maintenir le droit général des titulaires d'une maturité suisse à l'admission aux universités. Ce maintien doit être examiné également sous l'angle du deuxième objectif du gymnase, la maturité sociale.

D'un point de vue normatif, le chef de projet juge pertinent le double but actuel de la maturité, qui est de conférer une aptitude générale aux études supérieures et de faire acquérir une culture générale étendue, mais en inversant l'ordre d'importance. En effet, notre société aura à résoudre toute une série de problèmes importants à l'avenir. Or, dans notre culture, on choisit les problèmes à régler et la manière de les régler par la voie démocratique. Les élèves qui font une scolarité gymnasiale sont en majorité les plus performants sur le plan cognitif. Beaucoup de ces citoyennes et citoyens en herbe sont ainsi appelés à occuper des positions de responsabilité et à participer aux décisions. Le but d'acquérir des compétences permettant d'assumer des responsabilités dans la société est donc particulièrement adapté à leur cas. Si les tâches et les problèmes complexes ont souvent un caractère transversal, leurs différentes composantes sont *in fine* toujours spécifiques à une branche. De plus, il règne aujourd'hui dans les sciences de l'éducation un large consensus autour de la notion – qui se démarque des théories formelles de l'éducation déjà anciennes et de leurs variantes plus récentes (comme la théorie des qualifications clés) – que la résolution de problèmes, qu'ils soient transversaux ou spécifiques à une branche, requiert la possession de connaissances et de compétences dans les domaines concernés. La connaissance de méthodes générales, par exemple, est tout aussi insuffisante que la formation dans un domaine non concerné. C'est la raison pour laquelle les contenus de l'enseignement destinés à conférer une maturité permettant d'assumer des responsabilités sociales doivent être choisis non pas de manière arbitraire, mais avec soin et en se rapportant au but visé. Leur appropriation s'inscrit naturellement dans un processus d'apprentissage actif qui prend du temps. Des problèmes de société se posent dans de nombreux domaines; il faut donc offrir une éducation dans des domaines variés. L'abandon du but de conférer une aptitude générale aux études supérieures au profit d'une maturité spécialisée est en conflit avec le but de développer une maturité sociale à travers une vaste culture générale.

Le maintien du droit général à l'admission dans toutes les branches d'études et la diversité des disciplines gymnasiales qu'il implique sont en harmonie avec le but de la maturité sociale. Reste à résoudre le problème du choix précis des disciplines et de leur curriculum compte tenu du temps limité à disposition pour l'enseignement au gymnase. Les études réalisées dans le cadre d'EVAMAR II confirment des constats antérieurs (par ex. Huber, 1998 et réédition en cours d'impression) selon lesquels une aptitude générale aux études supérieures suppose dans tous les cas la possession de compétences de base en langue première et en mathématiques, mais aussi – comme le montre l'enquête auprès des professeurs – en anglais en sa qualité de langue de la science. Mais ces compétences d'études élémentaires ne représentent qu'une partie des curricula de ces disciplines et des examens correspondants, comme EVAMAR II l'a montré en particulier pour la langue première. Les curricula gymnasiaux portent sur des domaines plus ou moins étendus dont l'enseignement est justifié par le but de l'acquisition d'une maturité sociale. Dans quelle mesure les curricula gymnasiaux contiennent-ils des éléments qui ne sont pas nécessaires au regard de leurs deux objectifs (aptitude générale aux études et maturité sociale)?

La question reste ouverte: elle ne fait pas partie du mandat d'EVAMAR II. Les analyses conduites dans le cadre des sous-projets A1, A3 et A4 ont en outre montré que la plupart des branches d'études analysées requièrent aussi des connaissances et des compétences spécifiques à leur domaine. Mais si toutes les disciplines d'études requéraient des connaissances de départ spécialisées dans leur domaine, le gymnase serait vite débordé. Selon l'art. 5 du RRM, les écoles de maturité n'ont d'ailleurs pas à répondre à cette exigence: il leur est expressément demandé d'éviter la spécialisation. Les lignes directrices du Plan d'études cadre précisent d'ailleurs que le gymnase «ne doit pas dispenser une spécialisation propédeutique pour certaines filières» (CDIP, 1994, p. 6). Le gymnase est donc tenu de se limiter aux connaissances «fondamentales» exigées à l'entrée des différentes filières d'études. Cela s'applique aussi aux disciplines universitaires dont il est demandé, dans le cadre du débat actuel sur la politique de l'éducation, que les disciplines gymnasiales qui les préfigurent soient renforcées au nom de l'amélioration de la jonction avec l'université. Quelles sont ces connaissances fondamentales dans les différentes disciplines? Cette question reste sans réponse fondée. Un renvoi à la tradition en place nous semble un argument insuffisant. Il en va de même de la déduction didactique naïve qui voudrait que, si une branche d'études revêt un niveau élevé d'importance pour la société alors qu'elle affiche un taux de fréquentation faible, il faut renforcer l'enseignement gymnasial correspondant. On peut donc se demander si le gymnase ne doit pas chercher *au premier chef* à former des personnes capables d'assumer des responsabilités dans la société. Le maintien d'un large éventail de disciplines dans l'enseignement gymnasial serait alors assuré et les universités pourraient s'appuyer sur les compétences spécifiques aux disciplines dont l'acquisition est requise pour atteindre cet objectif de formation. De plus, les branches universitaires ne seraient pas en concurrence pour obtenir des heures dans l'enseignement gymnasial.

4 Assurer une compensation à l'entrée dans les filières d'études

A notre avis, il n'est pas possible dans la pratique que la maturité continue de conférer un droit général à l'admission aux universités, comme nous le préconisons, sans que certains bacheliers et bachelères aient, de cas en cas, des lacunes dans certaines disciplines. Le problème inévitable que pose le but utopique d'arriver à une «aptitude sans lacune aux études» doit donc être atténué par une extension de l'offre de cours de compensation organisés par les universités dans leur domaine d'études. Les universités proposent déjà des cours pour combler des lacunes en mathématiques, par exemple. Si la recommandation formulée à la section 2 est mise en œuvre, il ne sera plus nécessaire de proposer des cours de compensation en mathématiques et dans les autres disciplines pour lesquelles des standards de base ont été définis et dont les notes ne sont plus compensables. En outre, lorsque les responsables des différentes branches d'études définissent les connaissances requises pour rentrer dans leur filière, ils devraient tenir compte du fait que la mission du gymnase n'est pas de préparer les élèves à une branche d'études déterminée (lire aussi la section 3).

5 Développer la capacité d'apprendre de manière autonome sans restreindre l'éventail des disciplines

Les résultats du sondage auprès des professeurs et du questionnaire complémentaire soumis aux bachelères et aux bacheliers corroborent les études d'évaluation antérieures dans la mesure où la capacité d'apprendre et de travailler de manière autonome est décrite comme partiellement lacunaire dans les autoévaluations (Ramseier et al., 2005, Notter & Arnold, 2003 et 2006). Il convient donc de se rallier à ces études antérieures: un poids plus grand doit être donné à l'encouragement ciblé de la capacité d'apprendre de manière autonome dans les gymnases. Cela aurait pour effet non seulement de développer chez les bachelères et les bacheliers des compétences générales essentielles pour les études supérieures, mais aussi de leur donner de meilleurs outils pour combler eux-mêmes les lacunes qu'ils pourraient avoir en fonction de la filière choisie, lacunes inévitables comme nous l'avons vu. Cependant, il ne faut pas que le développement de la capacité d'apprendre et de

travailler de manière autonome soit renforcé au détriment du contenu ou de la variété des disciplines gymnasiales, au risque de créer des lacunes spécifiques par rapport au large éventail des études universitaires et de compromettre la réalisation de l'autre objectif du RRM (situé hors du champ d'étude d'EVAMAR II), soit apporter aux élèves une maturité sociale fondée sur une large culture générale. Cela suppose de se concentrer encore mieux sur l'essentiel dans toutes les disciplines, par exemple en se fondant sur une meilleure analyse des nécessités compte tenu de l'objectif d'apporter une maturité sociale et en améliorant la concertation avec les hautes écoles au sujet des attentes à l'entrée dans le cycle tertiaire.

Il ne s'agit pas de recommander la création d'une discipline «Apprendre à apprendre de manière autonome». La pratique scolaire fournit suffisamment d'exemples de tentatives infructueuses (défiance du transfert des compétences dans les disciplines, méthodes inadaptées aux disciplines, cessation prématurée du guidage des processus d'apprentissage nécessaires pour permettre aux élèves de passer le cap de l'internalisation de ces processus). Pour que la capacité d'apprendre de manière autonome s'intègre dans l'enseignement, elle doit constituer un axe de la didactique de l'enseignement des disciplines. Ce postulat a des conséquences de grande portée, qui peuvent s'étendre à la formation et au perfectionnement du corps enseignant des gymnases.

6 Améliorer la maîtrise de la langue première

Combler les déficits de maîtrise de la langue première constatés chez une partie des bacheliers et des bacheliers n'incombe pas seulement à la discipline gymnasiale correspondante. En effet, les lacunes observées se rapportent à l'utilisation de la langue première, qui est pratiquée dans la plupart des autres disciplines gymnasiales et qui est une compétence requise bien avant le stade des études universitaires. La maîtrise de la langue première peut donc être améliorée par un encouragement systématique à utiliser correctement cette langue dans toutes les disciplines. Cet aspect doit aussi être abordé dans le cadre de la formation et du perfectionnement des membres du corps enseignant du gymnase, quelle que soit leur discipline.

7 Faire davantage référence à des standards dans les programmes d'enseignement

Les grandes différences observées dans les résultats au test non seulement entre des individus, mais aussi entre des classes entières, ne peuvent pas s'expliquer uniquement par des différences de réalisation des objectifs; elles sont dues également à des différences dans les curricula effectivement enseignés. Le Plan cadre d'études suisse (PEC) n'a pas un caractère suffisamment obligatoire. Les plans d'études des cantons ou des écoles peuvent s'en écarter de manière substantielle et ils contribuent relativement peu à la prescription de normes de qualité dans certains domaines. C'est très certainement l'une des raisons pour lesquelles les examens de maturité analysés présentent des niveaux d'exigence très disparates. Si les examens présentant un niveau d'exigence élevé et les résultats aux tests satisfaisants à excellents, majoritaires les uns et les autres, ne requièrent pas de consignes plus contraignantes, il est cependant recommandé d'établir des standards minimaux ayant un caractère obligatoire. Ils permettraient d'éviter que les résultats ne soient tirés vers le bas, comme cela a été observé, sans pour autant brider la recherche de l'excellence. En outre, l'existence de standards minimaux obligatoires ne porterait pas atteinte à la diversité des contenus dans la partie supérieure des exigences.

8 Faire davantage référence à des standards dans l'évaluation des performances

Les différences observées entre les classes sont en partie très importantes. Les notes vont en gros dans la même direction que les résultats du test, mais il y a aussi des «anomalies». Les enseignantes et les enseignants du gymnase manient visiblement très bien les standards de qualité interclasses, mais il y a lieu de supposer que ceux-ci ont été adaptés au potentiel de certaines classes. Cela diminue la valeur significative des notes de maturité. Il convient donc de réfléchir à la manière dont on peut renforcer la référence à des standards. La centralisation des examens de maturité est plutôt à déconseiller car elle présente de multiples inconvénients, largement exposés dans la littérature spécialisée. Mais il faudrait au moins permettre et encourager une évaluation objectivée des résultats obtenus par référence à des standards minimaux. A cet effet, on peut recommander les mesures suivantes:

- La référence à des standards dans l'enseignement et les examens doit occuper une place plus importante dans la formation et le perfectionnement du corps enseignant.
- Il faut étudier des solutions intermédiaires pour trouver un dispositif standardisé de mesure des résultats qui soit moins centralisé qu'une maturité unique mais qui permette néanmoins une référence efficace à des standards inter-écoles. On peut notamment employer la méthode consistant mettre à la disposition des membres du corps enseignant un système de tests construits sur la base de modèles de compétence, auxquels ils ont accès à l'issue d'une période d'enseignement d'une durée déterminée (entre un trimestre et deux ans) et grâce auxquels ils peuvent évaluer le degré de divergence entre les standards et les objectifs atteints par leurs élèves. Cette forme d'évaluation serait volontaire et les informations seraient communiquées uniquement aux enseignantes et aux enseignants. Les expériences de ce type conduites à l'école obligatoire (système «Klassencockpit» avec modules d'orientation dans le canton de Saint-Gall) donnent à penser que ce genre de système peut avoir un effet autorégulateur positif.
- Les écoles devraient pouvoir évaluer leur performance individuelle sur la base du taux de réussite universitaire de leurs bachelières et bacheliers. Mais le calcul de ce critère de réussite suppose d'exploiter la statistique universitaire. Les systèmes d'information nécessaires sont en voie de développement dans la plupart des universités et dans les EPF. Il devrait donc être possible d'accéder aux données suivantes, qui sont utiles entre autres pour le monitoring du gymnase:
 - Saisie du choix du profil de formation au gymnase (options spécifiques et options complémentaires, domaine thématique du travail de maturité) et des notes de maturité sur la base des certificats de maturité présentés et établissement d'un lien avec le choix de la filière d'études et la réussite dans cette filière (notes aux examens et autres attestations de niveau).
 - Relevé systématique des parcours universitaires et de toutes leurs étapes (nombre de changements d'orientation, interruptions, durée des études, types de diplôme [bachelor, master, doctorat, etc.]) et établissement d'un lien avec les données des diplômes.
 - Enregistrement des motifs ayant déterminé le choix de la branche et du lieu d'études au moyen d'un questionnaire d'entrée écrit standardisé (par ex. intérêt, bonnes notes au gymnase, parcours, langue de l'université, coût, etc.).

9 Maintenir le gymnase en quatre ans

Les résultats d'EVAMAR II indiquent que les gymnases où la durée minimale de la scolarité est de trois ans seulement sont désavantagés. Manifestement, les gymnases offrant une formation en trois ans associée aux degrés pré-gymnasiaux des écoles secondaires locales parviennent moins bien que les gymnases proposant un cursus en quatre ans minimum à développer chez leurs élèves des structures de connaissances et des compétences spécifiques et interdisciplinaires aussi bonnes pendant la même période. Les causes de ce

désavantage résident probablement dans le fait que les deux degrés sont conçus différemment, les conditions d'admission dans les gymnases en trois ans ne sont pas les mêmes et la structure de l'enseignement dans les différentes disciplines est probablement plus cohérente lorsqu'elle est assurée aux deux degrés par le corps enseignant gymnasial. Il convient de revoir les dispositions afférentes du RRM. Il faudrait au minimum qu'aucun nouveau canton ne puisse réduire la durée du gymnase à trois ans, même s'il s'agit de la catégorie d'établissement scolaire la plus onéreuse. Les données d'EVAMAR II montrent d'ailleurs que là où il existe un enseignement à caractère pré-gymnasial, qui allonge la durée du gymnase, le niveau final dans toutes les données collectées (y compris les notes de maturité) est supérieur, légèrement mais tout de même significativement.

10 Assurer l'acquisition d'une aptitude aux études supérieures dans toutes les options spécifiques

Les différences de niveau observées selon les groupes d'option spécifique sont parfois considérables. Elles apparaissent non seulement dans les tests d'EVAMAR II mais aussi, avec moins d'amplitude, dans les notes de maturité, qui sont attribuées par les écoles elles-mêmes. Même si cela n'a probablement pas d'effets pratiques handicapants du fait que les bacheliers et les bachelières, comme nous l'avons vu plus haut, «adaptent» le choix de leurs études, on peut tout de même penser que ces disparités produisent des phénomènes de canalisation qui posent problème au regard de l'objectif de conférer aux élèves une aptitude générale aux études. De plus, les compétences nécessaires pour un domaine d'études ne s'acquièrent pas exclusivement dans la discipline gymnasiale correspondante, loin s'en faut. Il est important pour réussir des études en sciences économiques, par exemple, d'avoir des compétences solides dans des disciplines autres que l'économie et le droit, notamment les mathématiques et la langue première.

Les différences de performances sont donc visibles dans les tests et dans les notes de maturité. Lorsque, dans les options spécifiques nouvelles, il arrive que plus de 50% des bacheliers et des bachelières obtiennent une note insuffisante à l'épreuve écrite de la maturité en mathématiques, on peut s'interroger sur leur aptitude générale à poursuivre des études supérieures. Apparemment, les notes attribuées lors des épreuves de maturité n'ont jamais encore fait l'objet d'une évaluation systématique à l'échelle de la Suisse.

Les résultats d'EVAMAR II ne confirment donc pas le constat fait par EVAMAR I (Ramseier et al., 2005, p. 138), et repris dans d'autres publications, que les bacheliers et les bachelières réalisent, dans les disciplines de base, des performances équivalentes dans le nouveau système d'options spécifiques et dans l'ancien système des maturités de type fixe. Les résultats observés dans le nouveau système sont en partie moins bons. L'étude EVAMAR II permet de corriger les constats antérieurs. Il serait cependant erroné, ou tout au moins prématuré, d'expliquer les hausses et les baisses de performances par l'impact de l'enseignement dans les options spécifiques concernées. Il est possible que des disparités existent déjà au moment du choix de l'option spécifique, et la dynamique d'interaction entre l'enseignement, l'apprentissage et l'évaluation qui se développe tout au long du gymnase peut être très différente selon les classes. Il faut néanmoins envisager des mesures pour améliorer les résultats qu'une partie des bacheliers et des bachelières obtiennent dans les parties de l'apprentissage des mathématiques et de la langue première qui sont pertinentes pour les études supérieures; cela concerne surtout ceux qui choisissent une option spécifique introduite par le RRM 95 ne correspondant pas à un ancien type de maturité. La suppression de la possibilité de compenser les notes insuffisantes dans des domaines partiels de base proposées à la section 2 (compétences de base) en est une. Il n'y aurait pas lieu de craindre une diminution du taux de réussite de 25% ou plus. Au contraire, une telle mesure aurait avant tout, à notre avis, un effet de guidage des efforts d'apprentissage tout au long de la formation gymnasiale.

11 Réduire les disparités entre les taux de maturité

L'hypothèse que des taux de maturité élevés s'accompagnent de résultats plutôt moins bons a été corroborée non seulement par les tests mais aussi, en grande partie, par le relevé des notes de maturité. Dans les cantons où les taux de maturité sont élevés, les taux d'admission au gymnase sont également plus importants. Comme rien ne permet de dire que les performances scolaires se développent moins dans les cantons ayant des taux de maturité élevés que dans les cantons ayant des taux de maturité bas, il y a tout lieu de supposer que la performance scolaire moyenne dans les premiers est inférieure déjà au stade de l'entrée au gymnase et qu'à niveau scolaire égal les élèves ont plus de chances d'accéder au gymnase dans les cantons où les taux de maturité sont élevés. De plus, il apparaît que la formation gymnasiale ne parvient pas à remédier à ces disparités. Cette inégalité des chances manifeste doit donner lieu à une discussion politique.

12 Poursuivre les efforts dans le domaine de l'égalité des sexes

Les différences entre les sexes suivent en partie les stéréotypes bien connus. Il faut donc poursuivre les efforts faits pour améliorer les performances des gymnasiens en langue première et des gymnasiennes en mathématiques et en sciences expérimentales ou pour accroître leur intérêt pour ces disciplines, dont on suppose qu'il explique leurs résultats. Il faudrait également analyser les disparités qui pourraient exister dans d'autres disciplines sur lesquelles ne portait pas EVAMAR II.

13 Développer le travail de maturité

L'introduction du travail de maturité apparaît comme un point positif. Le travail individuel sur des projets, qui est un élément important de l'apprentissage gymnasial, doit être développé de manière modérée en tenant compte des facteurs de réussite ressortant des enquêtes d'EVAMAR II.

14 Poursuivre les analyses

Le projet EVAMAR II est conçu comme un projet transversal. Les résultats obtenus soulèvent donc plus de questions qu'ils ne peuvent apporter de réponses fondées sur les phénomènes observés. Il importe donc d'étudier les questions restées sans réponse en pratiquant des analyses longitudinales et en exploitant systématiquement toutes les données personnelles générées au cours de la formation gymnasiale et des études supérieures qui s'ensuivent.

Bibliographie (sans le chapitre VI)

- Abraham, U., Beisbart, O., Koß, G., & Marenbach, D. (2003). *Praxis des Deutschunterrichts. Arbeitsfelder, Tätigkeiten, Methoden*. Donauwörth: Auer.
- Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research*, 58 (4), 375-404.
- Alexander, P. A., Pate, P. E., Kulikowich, J. M., Farrell, D. M., & Wright, N. L. (1989). Domain specific and strategic knowledge: Effects of training on students of differing ages or competence levels. *Learning and Individual Differences*, 1 (3), 283-325.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80, 260-267.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Raths, J., et al. (2001). *A taxonomy for learning, teaching and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Artelt, C., & Schlagmüller, M. (2004). Der Umgang mit literarischen Texten als Teilkompetenz im Lesen? Dimensionsanalysen und Ländervergleiche. Dans: U. Schiefele, C. Artelt, W. Schneider & P. Stanat (éd.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (p. 169-196). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Artelt, C., Stanat, P., Schneider, W., Schiefele, U., & Lehmann, R. (2004). Die PISA-Studie zur Lesekompetenz: Überblick und weiterführende Analysen. Dans: U. Schiefele, C. Artelt, W. Schneider & P. Stanat (éd.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (p. 139-168). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart, and Winston.
- Bamberger, R., & Vanecek, E. (1984). *Lesen-Verstehen-Lernen-Schreiben. Die Schwierigkeitsstufen von Texten in deutscher Sprache*. Wien: Jugend und Volk, Diesterweg, Sauerländer.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Baumert, J., Bos, W., & Lehmann, R. (éd.) (2000). *TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*. Opladen: Leske + Budrich.
- Baumert, J., Brunner, M., Lüdtke, O., & Trautwein, U. (2007). Was messen internationale Schulleistungstudien? – Resultate kumulativer Wissenserwerbsprozesse. *Psychologische Rundschau*, 58 (2), 118-145.
- Beck, B., & Klieme, E. (éd.) (2007). *Sprachliche Kompetenzen: Konzepte und Messung: DESI-Studie (Deutsch Englisch Schülerleistungen International)*. Weinheim: Beltz.
- Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly*, 26, 251-276.

- Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders*, 25, 62-80.
- Biologie heute entdecken (2004). *Allgemeine Ausgabe für SII*. Braunschweig: Schroedel.
- Biologie Oberstufe (2001). *Gesamtband*. Berlin: Cornelsen.
- Björnsson, C. H. (1968). *Läsbarhet* (version allemande: *Lesbarkeit durch Lix, Übersetzung von Studienrätin Helga Johansson, 1968*). Stockholm: Pedagogiskt Centrum vid Stockholms skoldirektion.
- Bloom, B. S. (1974). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim et Bâle: Beltz.
- Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay.
- Bol, L., & Strage, A. (1996). The contradiction between teachers' instructional goals and their assessment practices in high school biology courses. *Science Education*, 80 (2), 145-163.
- Bonfadelli, H., Jarren, O., & Siegert, G. (éd.) (2005). *Einführung in die Publizistikwissenschaft*. Berne: Haupt.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. überarbeitete Auflage). Berlin: Springer.
- Boueke, D. (éd.) (1979). *Deutschunterricht in der Diskussion: Forschungsberichte* (2. Auflage). Paderborn: Schöningh.
- Britton, B. K., Gulgoz, S., & Glynn, S. (1993). Impact of good and poor writing on learners: Research and theory. Dans: B. K. Britton, A. Woodward & M. R. Binkley (éd.), *Learning from textbooks: Theory and practice* (p. 1-46). Hillsdale: Erlbaum.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (2. Auflage). Munich: Pearson Studium.
- Bybee, R. (1997). *Achieving scientific literacy: from purposes to practices*. Portsmouth: Heilman.
- Campbell, N. A., & Reece, J. B. (2006). *Biology* (6th edition). Munich: Pearson Education.
- Carlsen, W. S. (1991). Questioning in classrooms: A sociolinguistic perspective. *Review of Educational Research*, 61, 157-178.
- Cazden, C. B. (1986). Classroom discourse. Dans: M. C. Wittrock (éd.), *Handbook of research on teaching* (p. 432-463). New York: Macmillan.
- CDIP (1994). *Plan d'études cadre pour les écoles de maturité (Dossier 30B)*. Berne: Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP).
- CDIP (1995). *Ordonnance du Conseil fédéral/Règlement de la CDIP des 16 janvier/15 février 1995 sur la reconnaissance des certificats de maturité gymnasiale (RRM)*. Berne: Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP).
- CDIP (2004a). HarmoS - Finalités et conception du projet. Berne: CDIP.
http://www.edudoc.ch/static/web/arbeiten/harmos/weissbuch_f.pdf. Consulté le 20.6.2011.
- CDIP (2004b). Conférence de presse annuelle de la CDIP: harmonisation des objectifs de l'école obligatoire. Communiqué de presse du 1.7.2004. Berne: CDIP.
<http://www.edk.ch/dyn/14013.php>. Consulté le 20.6.2011.

- CDIP (2005). HarmoS. Composition des consortiums chargés du développement scientifique. http://www.edudoc.ch/static/web/arbeiten/harmos/harmos_kons_f.pdf. Consulté le 20.6.2011.
- CDIP (2007a). Accord intercantonal du 14 juin 2007 sur l'harmonisation de la scolarité obligatoire (concordat HarmoS). Berne: CDIP. http://edudoc.ch/record/24710/files/HarmoS_f.pdf. Consulté le 20.6.2011.
- CDIP (2008). Accord intercantonal du 14 juin 2007 sur l'harmonisation de la scolarité obligatoire (concordat HarmoS): état d'avancement des procédures cantonales d'adhésion. Berne: CDIP. http://www.edudoc.ch/static/web/arbeiten/harmos/liste_rat_df.pdf. Consulté le 20.6.2011.
- Chamorro-Premuzic, T., & Furnham, A. (2004). A possible model for explaining the personality-intelligence interface. *British Journal of Psychology*, 95, 249-264.
- Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence, approaches to learning and academic performance. *Personality and Individual Differences*, 44, 1596-1603.
- Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning & Verbal Behavior*, 18, 257-273.
- Chipman, S. F., Segal, J. W., & Glaser, R. (éd.) (1985). *Thinking and learning skills, vol.2*. Hillsdale, N.J.: Erlbaum.
- Cochran, W. G. (1977). *Sampling Techniques*. New York: John Wiley & Sons.
- Commission des bibliothèques scolaires et des bibliothèques communales du canton de Berne (éd.) (1989). Textbeurteilung: eine vielschichtige Angelegenheit. Lesbarkeits-Index (LIX): ein wertvolles Hilfsmittel. *Der Berner Bibliothekar*, 46, 2-12.
- Commission suisse de maturité (CSM) (2008). Directives pour l'examen suisse de maturité 2003-2006 (prolongées jusqu'en 2008). Domaine des sciences expérimentales. Berne: Secrétariat d'Etat à l'éducation et à la recherche SER. [Ancienne version non disponible. Version actuelle: http://www.sbf.admin.ch/htm/themen/bildung/matur/ch-matur_fr.html. Consulté le 20.6.2011.]
- Conseil de l'Europe (2001). Un cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer. Division des politiques linguistiques, Strasbourg. http://www.coe.int/t/dg4/linguistic/Source/Framework_FR.pdf. Consulté le 21.3.2011.
- Conseil fédéral (1968). Ordonnance du 22 mai 1968 sur la reconnaissance de certificats de maturité. Recueil officiel du droit fédéral: RO 1968 717.
- Conseil fédéral (2007). Ordonnance du 15 février 1995 sur la reconnaissance des certificats de maturité gymnasiale (ORM). Modification du 27 juin 2007. Recueil officiel du droit fédéral: RO 2007 3477.
- DESI-Konsortium (2006). Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch-Englisch-Schülerleistungen-International (DESI). Frankfurt-sur-le-Main: Deutsches Institut für Internationale Pädagogische Forschung (DIPF). http://www.dipf.de/desi/DESI_Zentrale_Befunde.pdf. Consulté le 14.3.2006.
- DESI-Konsortium (éd.) (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- Diakidoy, I. N. (1999). Comprehension and learning from scientific text. Dans: A. Gagatsis (éd.), *A multidimensional approach to learning in mathematics and science* (p. 385-413). Nicosie, Chypre: Intercollege Press.

- Dochy, F. J. R. C. (1992). *Assessment of prior knowledge as a determinant for future learning: The use of prior knowledge state tests and knowledge profiles*. Londres: Lemma B.V.
- Dochy, F. J. R. C., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research*, 69, 145-186.
- Dresel, M., Stöger, H., & Ziegler, A. (2006). Klassen- und Schulunterschiede im Ausmaß von Geschlechtsunterschieden bei Leistungsbewertungen und Leistungsaspirationen: Ergebnisse einer Mehrebenenanalyse. *Psychologie in Erziehung und Unterricht*, 53, 44-61.
- Drinkmann, A., & Groeben, N. (1989). *Metaanalysen für Textwirkungsforschung*. Weinheim: Deutscher Studienverlag.
- Eberle, F. (1986). *Unterschiede in schulleistungsrelevanten Merkmalen zwischen Wirtschaftsgymnasiasten und Gymnasiasten anderer Maturitätstypen*. Bamberg: Difo-Druck, 2 volumes.
- Eberle, F. (2005). *Definitives Konzept für die Phase II des Projekts «Evaluation der Maturitätsreform MAR 95» (EVAMAR) vom 11. April 2005 zuhanden des Bundesamts für Bildung und Wissenschaft und der Konferenz der kantonalen Erziehungsdirektoren (EDK)* (document interne). Zurich: Université de Zurich.
- Eberle, F. (2008). Ist das Bildungsziel des MAR noch zeitgemäss? *AMV aktuell*, 2, 19-28.
- Ehlich, K. (2003). Universitäre Textarten, universitäre Struktur. Dans: K. Ehlich & A. Steets (éd.), *Wissenschaftlich schreiben – lehren und lernen* (p. 13-28). Berlin: De Gruyter.
- Ehlich, K., & Steets, A. (2003). Wissenschaftliche Schreibanforderungen in den Disziplinen. Eine Umfrage unter ProfessorInnen der LMU. Dans: K. Ehlich & A. Steets (éd.), *Wissenschaftlich schreiben – lehren und lernen* (p. 129-154). Berlin: De Gruyter.
- Elliot, A. J., & Thrash, T. M. (2001). Achievement goals and the hierarchical model of achievement motivation. *Educational Psychology Review*, 13, 139-156.
- Ernst, R. (2004). Verantwortungsbewusste Bildung für die Zukunft: Eine Herausforderung für Mittel- und Hochschule. *Gymnasium Helveticum*, 1, 6-10.
- Feilke, H., & Steinhoff, T. (2003). Zur Modellierung der Entwicklung wissenschaftlicher Schreibfähigkeiten. Dans: K. Ehlich & A. Steets (éd.), *Wissenschaftlich schreiben – lehren und lernen* (p. 112-128). Berlin: De Gruyter.
- Fenollar, P., Roman, S., & Cuestas, P. J. (2007). University students' academic performance: An integrative conceptual framework and empirical analysis. *British Journal of Educational Psychology*, 77 (4), 873-891.
- Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien – ein Problemaufriss. Dans: H. Mandl & H. F. Friedrich (éd.), *Lern- und Denkstrategien. Analyse und Intervention*. (p. 3-54). Gottingue: Hogrefe.
- Früh, W. (2004). *Inhaltsanalyse. Theorie und Praxis*. Stuttgart: UTB.
- Fucks, W. (1955). *Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen*. Cologne: Westdeutscher Verlag.
- Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, 14, 49-66.
- Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.

- Georghiades, P. (2000). Beyond conceptual change learning in science education: focusing on transfer, durability and metacognition. *Educational Research*, 42 (2), 119-139.
- Glinz, E., Glinz, H., & Ramseier, M. (1997). *Sprachunterricht – Theorie und Praxis: Grundlagen zum «Schweizer Sprachbuch»*. Zurich: SABE.
- Glinz, H. (1987). Die Sprachtheorie in und hinter den Lehrern und die Entwicklung der Sprachfähigkeit in den Schülern. Dans: R. Wimmer (éd.), *Sprachtheorie. Der Sprachbegriff in Wissenschaft und Alltag. Jahrbuch des Instituts für deutsche Sprache (Sprache der Gegenwart 71)*. Düsseldorf: Schwann.
- Gräber, W., Nentwig, P., & Nicolson, P. (éd.) (2002). *Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung*. Opladen: Leske + Budrich.
- Groeben, N. (1972). *Die Verständlichkeit von Unterrichtstexten: Dimensionen und Kriterien rezeptiver Lernstadien*. Münster: Aschendorff.
- Hager, W. (1995). *Programme zur Förderung des Denkens bei Kindern*. Göttingen: Hogrefe.
- Haider, G., & Reiter, C. (éd.) (2001). *Pisa 2000: Nationaler Bericht. Internationale und nationale Ergebnisse. Vergleich der Schülerleistungen mit dem Schwerpunkt Lesen und Leseverständnis. Beiträge zur vergleichenden Schulforschung*. Innsbruck: Studien-Verlag.
- Haider, G., & Reiter, C. (éd.) (2004). *Pisa 2003 – Internationaler Vergleich von Schülerleistungen. Ergebnisse im Überblick. Executive Summary*. Graz: Universität de Salzburg, division des sciences de l'éducation.
- Hailikari, T., Nevgi, A., & Komulainen, E. (2008). Academic self-beliefs and prior knowledge as predictors of student achievement in mathematics: A structural model. *Educational Psychology Review*, 28 (1), 59-71.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15 (3), 309-334.
- Halpern, D. F., & LaMay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review*, 12, 229-246.
- Hänsgen, K.-D., & Spicher, B. (2001). *EMS Eignungstest für das Medizinstudium in der Schweiz. Vorhersage des Prüfungserfolges*. Fribourg: Université de Fribourg.
- Hänsgen, K.-D., & Spicher, B. (2007). *EMS Eignungstest für das Medizinstudium in der Schweiz 2007*. Fribourg: Université de Fribourg.
- Heller, K. A. (éd.) (2002). *Begabtenförderung im Gymnasium. Ergebnisse einer zehnjährigen Längsschnittstudie*. Opladen: Leske + Budrich.
- Hornung, A. (2005). Textkompetenz und Sprachkompetenz – einige praxisgeleitete Überlegungen zu einem aktuellen Thema. Dans: Höheres Lehramt Berufsschulen und Höheres Lehramt Mittelschulen am Zürcher Hochschulinstitut für Schulpädagogik und Fachdidaktik (éd.), *Texte verstehen – Berichte aus Forschung und Praxis Bd. 2* (p. 128-135). Berne et Zurich: hep/Pestalozzianum.
- Huber, L. (1997). Fähigkeit zum Studieren – Bildung durch Wissenschaft. Zum Problem der Passung zwischen Gymnasialer Oberstufe und Hochschule. Dans: E. Liebau, W. Mack & C. Scheilke (éd.), *Das Gymnasium. Alltag, Reform, Geschichte, Theorie* (p. 333-351). Weinheim: Juventa.
- Huber, L. (1998). Allgemeine Studierfähigkeit, basale Fähigkeiten, Grundbildung. Zur aktuellen Diskussion um die gymnasiale Oberstufe. Dans: R. Messner, E. Wicke & D. Bosse (éd.), *Die Zukunft der gymnasialen Oberstufe* (p. 150-181). Weinheim: Beltz.

- Huber, L. (2009). Von «basalen Fähigkeiten» bis «vertiefte Allgemeinbildung»: Was sollen Abiturienten für das Studium mitbringen? Dans: D. Bosse (éd.), *Gymnasiale Bildung zwischen Kompetenzorientierung und Kulturarbeit*. Wiesbaden: Verlag für Sozialwissenschaften.
- Ingendahl, W. (1999). *Sprachreflexion statt Grammatik. Ein didaktisches Konzept für alle Schulstufen*. Tübingen: Max Niemeyer.
- Jordan, A., Ross, N., Krauss, S., Baumert, J., Blum, W., Neubrand, M., et al. (2006). *Klassifikationsschema für Mathematikaufgaben: Dokumentation der Aufgabenklassifikation im COACTIV-Projekt. Materialien aus der Bildungsforschung, Nr. 81*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, 35, 1567-1577.
- Kim, S., & Van Dusen, L. M. (1998). The role of prior knowledge and elaboration in text comprehension and memory: A comparison of self-generated and text provided elaboration. *American Journal of Psychology*, 111, 353-378.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163-182.
- Kintsch, W. (1998). *Comprehension: a paradigm for cognition*. Cambridge: Cambridge University Press.
- Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.
- Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik*, 52, 876-903.
- Klieme, E., Neubrand, M., & Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse. Dans: J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiss (éd.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Klieme, E., Neubrand, M., & Lüdtke, O. (2002). *Kompetenzstufen und Schwierigkeitsmodelle für den PISA-Test zur mathematischen Grundbildung*. Opladen: Leske + Budrich.
- KMK (2004). Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland: Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie (Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 05.02.2004). http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Biologie.pdf. Consulté le 20.6.2011.
- KMK (2005). Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland: Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss (Jahrgangsstufe 10) (Beschluss der Kultusministerkonferenz vom 16.12.2004). Munich: Wolters Kluwer. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Biologie.pdf. Consulté le 20.6.2011.
- Köller, O., & Klieme, E. (2000). Geschlechtsdifferenzen in den mathematisch-naturwissenschaftlichen Leistungen. Dans: J. Baumert, W. Bos & R. Lehmann (éd.), *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 2. Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe* (p. 373-404). Opladen: Leske + Budrich.

- Köller, O., Watermann, R., Trautwein, U., & Lüdtke, O. (2004). *Wege zur Hochschulreife in Baden-Württemberg – TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien*. Opladen: Leske + Budrich.
- Köster, J. (2005). Wodurch wird ein Text schwierig? Ein Test für die Fachkonferenz. *Deutschunterricht*, 5, 34-39.
- Kretschmer, I. (1981). Untersuchung zur Förderung des produktiven Denkens von elfjährigen Schülern. *Die deutsche Schule*, 73 (4), 205-214.
- Kruse, O. (2007). Schreibkompetenz und Studierfähigkeit. Mit welchen Schreibkompetenzen sollten die Schulen ihre Absolvent/innen ins Studium entlassen? Dans: M. Becker-Mrotzek & K. Schindler (éd.), *Texte schreiben. Kölner Beiträge zur Sprachdidaktik* (5) (p. 117-143). Duisbourg: Gilles & Francke.
- Kubinger, K. D. (2003). Adaptives Testen. Dans: K. D. Kubinger & R. S. Jäger (éd.), *Schlüsselbegriffe der Psychologischen Diagnostik* (p. 1-9). Weinheim: Beltz.
- Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., et al. (2002). *Pisa 2000 – Dokumentation der Erhebungsinstrumente*. Berlin: MPI für Bildungsforschung.
- Lehtonen, R., & Pahkinen, E. J. (1994). *Practical Methods for Design and Analysis of Complex Surveys*. New York: John Wiley & Sons.
- Leutner, D., Wirth, J., & Fleischer, J. (2005). Zentrale Lernstandserhebungen in der Jahrgangsstufe 9 im Jahr 2004 in NRW: Erster Kurzbericht zur wissenschaftlichen Begleitung. Universität de Duisburg-Essen. Lehrstuhl für Lehr-Lernpsychologie.
http://www.standardsicherung.nrw.de/lernstand8/upload/download/ergebn_05/kurzbericht1_wb_04.pdf. Consulté le 20.6.2011.
- Levack, B. P. (2003). *Hexenjagd. Die Geschichte der Hexenverfolgungen in Europa*. Munich: Beck.
- Linder Biologie (2005). *Gesamtband SII* (22. Auflage). Brunswick: Schroedel.
- Linneweber, H., & Wälti, B. (2006). *Zwischenbericht HarmoS Konsortium Mathematik – für die erste Projektetappe: Mai 2005 – Juni 2006, Stand der Arbeiten August 2006*. Aarau: Fachhochschule Nordwestschweiz.
- Long, D. L., Seely, M. R., & Oppy, B. J. (1996). The availability of causal information during reading. *Discourse Processes*, 22, 145-170.
- Maradan, O., & Mangold, M. (2005). Bildungsstandards in der Schweiz. Das Projekt HarmoS. *PH-akzente*, 2, 3-7.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning II: Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology*, 46, 115-127.
- Mayr, E. (2001). *What Evolution Is*. New York: Basic Books.
- McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist*, 52, 509-516.
- McNamara, D. S. (2001). Reading both high-coherence and low coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, 55, 51-62.
- McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247-287.
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2005a). Kompetenzniveaus in den 2005 im Rahmen der Lernstandserhebungen 9 getesteten

Bereichen. Anforderungsniveaus der Aufgaben und Kompetenzniveaus der Schülerinnen und Schüler. Soest: Landesinstitut für Schule und Weiterbildung.
http://www.standardsicherung.nrw.de/lernstand8/upload/download/ergebn_05/kompetenzniveaus_05.pdf. Consulté le 20.6.2011.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2005b). Zentrale Lernstandserhebungen in Klasse 9 – Landesweite Ergebnisse 2005. Soest: Landesinstitut für Schule und Weiterbildung.
http://www.standardsicherung.nrw.de/lernstand8/upload/download/ergebn_05/lse-ergebnisse_2005.pdf. Consulté le 20.6.2011.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2007). Erläuterungen zu den Kompetenzniveaus Deutsch – Leseverstehen bei der Lernstandserhebung 2007 (Lernstand 8). Soest: Stab Bildungsforschung und Grundlagen der Standardentwicklung und -überprüfung (StB).
http://www.standardsicherung.schulministerium.nrw.de/lernstand8/upload/download/erg_06-07/komp-niv_de leseverstehen07.pdf. Consulté le 20.6.2011.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2008). Zentrale Lernstandserhebungen in der Jahrgangsstufe 8. Soest: Stab Bildungsforschung und Grundlagen der Standardentwicklung und -überprüfung (StB).
<http://www.standardsicherung.schulministerium.nrw.de/lernstand8/aktuelles/>. Consulté le 20.6.2011.

MNU (Deutscher Verein zur Förderung des mathematischen und naturwissenschaftlichen Unterrichts) (2005). *Arbeiten mit den Bildungsstandards im Fach Biologie fachspezifisch und fachübergreifend, dimensioniert und niveauvoll. Empfehlungen für die Umsetzung der KMK-Standards Biologie S I.*

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, 94, 249-259.

Mugabushaka, A.-M. (2005). Schlüsselqualifikationen im Hochschulbereich. Eine diskursanalytische Untersuchung der Modelle, Kontexte und Dimensionen in Deutschland und Großbritannien. Dissertation. Online-Publikation. Universität de Cassel.
http://kobra.bibliothek.uni-kassel.de/bitstream/urn:nbn:de:hebis:34-2118/1/dis3233_05.pdf. Consulté le 20.6.2011.

Narciss, S., & Proske, A. (2001). *Lern- und Studieraufgaben für die universitäre Lehre*. Technische Universität Dresden (non publié).

NCTM (2000). *National Council of Teachers of Mathematics: Principles and standards for school mathematics*. Reston, VA: NCTM.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.

Nold, G. (2007). DESI im Kontext des Gemeinsamen Europäischen Referenzrahmens für Sprachen. Dans: B. Beck & E. Klieme (éd.), *Sprachliche Kompetenzen: Konzepte und Messung: DESI-Studie (Deutsch Englisch Schülerleistungen International)* (p. 299-305). Weinheim: Beltz.

Notter, P., & Arnold, C. (2003). *Le passage aux études supérieures. Rapport relatif à un projet de la Conférence des directeurs de Gymnases suisses (CDGS) et de la Conférence des recteurs des universités suisses (CRUS)*. Berne: Secrétariat d'Etat à l'éducation et à la recherche SER (anc.: Office fédéral de l'éducation et de la science OFES)(= Dossiers OFES 2003, 5f).

- Notter, P., & Arnold, C. (2006). *Le passage aux études supérieures II. Rapport relatif à un projet de la Conférence des directeurs de Gymnases suisses (CDGS) et de la Conférence des recteurs des universités suisses (CRUS)*. Berne: Secrétariat d'Etat à l'éducation et à la recherche SER.
- O'Connor, M., & Paunonen, S. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, 43, 971-990.
- O'Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional «high-stakes» measures of high school students' science achievement. *American educational research journal*, 44 (1), 161-196.
- OCDE (éd.) (1999). *Mesurer les connaissances et compétences des élèves. Un nouveau cadre d'évaluation*. Paris: Organisation de coopération et de développement économiques (OCDE).
- OCDE (éd.) (2007). *PISA 2006: Les compétences en sciences, un atout pour réussir. Volume 1 - Analyse des résultats*. Paris: Organisation de coopération et de développement économiques (OCDE).
- Oelkers, J. (2006). *Bildungsstandards am Gymnasium: Ein neues Problem?* Conférence donnée lors du 2^e séminaire pour cadres de l'Institut *Sekundarstufe II* de la Haute école pédagogique du canton de Berne (PHBern), le 30 août 2006 à l'Université de Berne.
- Oelkers, J. (2008). *Die Qualität der Schweizer Gymnasien*. Berne: hep.
- Office fédéral de la statistique (OFS) (2008a). Examens finals des hautes écoles universitaires: Tableaux de base.
<http://www.bfs.admin.ch/bfs/portal/fr/index/themen/15/06/data.html>. Consulté le 17.6.2011.
- Office fédéral de la statistique (OFS) (2008b). Cartes interactives: Taux de maturité.
<http://www.bfs.admin.ch/bfs/portal/fr/index/themen/15/06/dos/blank/05/03.html>. Consulté le 17.6.2011.
- Ossner, J. (2006). Kompetenzen und Kompetenzmodelle im Deutschunterricht. *Didaktik Deutsch*, 21, 5-19.
- Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, 19, 228-242.
- Prenzel, M., Drechsel, B., Carstensen, C. H., & Ramm, G. (2004). PISA 2003 – eine Einführung. Dans: M. Prenzel, J. Baumert, W. Blum, R. H. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G. Rolff, J. Rost & U. Schiefele (éd.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (p. 13-46). Münster: Waxmann.
- Purves, W. K. (2008). *Life – The Science of Biology* (8th edition). Sunderland: Sinauer Associates.
- Ramseier, E., Allraum, J., Stalder, U., Grin, F., Alliata, R., Müller, S., et al. (2004). *Evaluation der Maturitätsreform 1995 (EVAMAR). Neue Fächerstruktur – Pädagogische Ziele – Schulentwicklung. Schlussbericht zur Phase 1*. Berne: Secrétariat d'Etat à l'éducation et à la recherche SER.
- Ramseier, E., Keller, C., & Moser, U. (1999). *Bilanz Bildung. Eine Evaluation am Ende der Sekundarstufe II auf der Grundlage der «Third International Mathematics and Science Study»*. Coire: Rüegger.
- Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, 80, 16-20.
- Resnick, L. (1987). *Education and learning to think*. Washington, D.C.: National Academy Press.
- Richards, E., & Singer, M. (2001). Representation of complex goal structures in narrative comprehension. *Discourse Processes*, 31, 111-135.
- Rost, J. (2004). *Lehrbuch Testtheorie Testkonstruktion*. Berne: Huber.

- Satin, A., & Shastry, W. (1999). *Survey Sampling*. Ottawa: Statistics Canada.
- Schallberger, U. (1996). Anforderungen an das Zulassungsverfahren in der Schweiz. Leitlinien für die Entwicklung eines eignungsdiagnostischen Verfahrens. Dans: K.-D. Hänsgen & N. Ischi (éd.), *Eignungsprüfung für das Medizinstudium. Kriterien und Testverfahren* (p. 92-101). Fribourg: Zentrum für Testentwicklung und Diagnostik.
- Schiefele, U., Artelt, C., Schneider, W., & Stanat, P. (éd.) (2004). *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schneider, G., & North, B. (1999). «In anderen Sprachen kann ich...» *Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit: Umsetzungsbericht. Nationales Forschungsprogramm 33 zur Wirksamkeit unserer Bildungssysteme*. Berne/Aarau: Centre suisse de coordination pour la recherche en éducation CSRE.
- Schüpbach, H., Pixner, J., & Zapf, S. (2006). Handlungskompetenz im Hochschulstudium. *Bildung und Erziehung*, 59 (2), 147-166.
- Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal*, 41 (1), 159-189.
- Sieber, P. (2006a). HarmoS Erstsprache. Conférence donnée lors du séminaire pour cadres sur les standards de formation du 27 janvier 2006 (non publiée).
- Sieber, P. (2006b). Schritte auf dem Weg zu einem Kompetenzmodell Erstsprache. Symposium Deutschdidaktik Weingarten. Zurich: Pädagogische Hochschule.
http://www.symposion-deutschdidaktik.de/fileadmin/template/download/beitraege/praesentationen_vortraege/11_Sieber.pdf. Consulté le 24.6.2011.
- Sieber, P. (2006c). Schritte auf dem Weg zu einem Kompetenzmodell Erstsprache. Texte mis en consultation en avril 2006. Document de travail non publié.
- Sieber, P. (éd.) (1994). *Sprachfähigkeiten – besser als ihr Ruf und nötiger denn je! : Ergebnisse und Folgerungen aus einem Forschungsprojekt*. Aarau: Sauerländer.
- Snow, C. P. (1963). *The Two Cultures: A Second Look*. Cambridge: Cambridge UP.
- Social Science Consulting. (2007). TextQuest – Software für Textanalysen. Osnabrück.
<http://www.textquest.de/pages/de/allgemeines.php>. Consulté le 24.6.2011.
- Stanat, P., & Kunter, M. (2001). Geschlechterunterschiede in Basiskompetenzen. Dans: Deutsches PISA-Konsortium (éd.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (p. 249-269). Opladen: Leske + Budrich.
- Steiner, G. (2001). *Lernen – 20 Szenarien aus dem Alltag. Eine Einführung in die Lernpsychologie* (3. Aufl.). Berne: Huber.
- Streblow, L. (2004). Zur Förderung der Lesekompetenz. Dans: U. Schiefele, C. Artelt, W. Schneider & P. Stanat (éd.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (p. 275-306). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Szidat, J. (2003). Erwartungen der Universität an die Mittelschule. *Bulletin der Vereinigung schweizerischer Hochschuldozenten*, 2/3, 14-17.
- Tait, H., & Entwistle, N. J. (1996). Identifying students at risk through ineffective study strategies. *Higher Education*, 31, 97-116.

- Takala, S. (2007). Relating Examinations to the Common European Framework. Dans: B. Beck & E. Klieme (éd.), *Sprachliche Kompetenzen: Konzepte und Messung: DESI-Studie (Deutsch Englisch Schülerleistungen International)* (p. 306-331). Weinheim: Beltz.
- Tauber, M. (1984). *Leserangepasste Verständlichkeit: der Einfluss von Lesbarkeit und Gliederung am Beispiel von Zeitungsartikeln*. Berne: Lang.
- Thim-Mabrey, C. (2005). Stilnormen als Textsortennormen. Korrektur und Beratung zu Texten von Schülern und Studierenden. Dans: K. Adamzik & W.-D. Krause (éd.), *Text-Arbeiten. Textsorten im fremd- und muttersprachlichen Unterricht an Schule und Hochschule* (p. 31-43). Tübingue: Gunter Narr.
- Thompson, R. A., & Zamboanga, B. L. (2004). Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology*, 96, 778-784.
- Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research*, 64 (1), 37-54.
- Trappmann, S., Hell, B., Weigand, S., & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs – eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 21 (1), 11-27.
- Trautwein, U., Watermann, R., Lüdtke, O., Köller, O., Baumert, J., Maaz, K., et al. (2006). *Transformation des Sekundarschulsystems und akademische Karrieren (TOSCA). Dokumentation der eingesetzten Items und Skalen*. Berlin: Max – Planck – Institut für Bildungsforschung.
- Trost, G. (1989). A nationwide testing program for admission to medical schools in West Germany. Dans: R. C. King & J. K. Collins (éd.), *Social applications and issues in psychology* (p. 131-137). Amsterdam: Elsevier Science.
- Tzeng, Y., van den Broek, P., Kendeou, P., & Lee, C. (2005). The computational implementation of the Landscape Model: Modeling inferential processes and memory representations of text comprehension. *Behavioral Research Methods*, 37, 277-286.
- van den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology*, 22, 335-351.
- van den Broek, P., Kendeou, P., Sung, Y. C., & Chen, M. (2003). *Comprehension and memory of science texts: A simulation using the Landscape Model*. Paper presented at the Annual Meeting of the Society for Text and Discourse. Madrid, June 2003.
- van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory- based and constructionist approaches in accounts of reading comprehension. *Discourse Processes*, 39, 299-316.
- van den Broek, P., Ridsen, K., Fletcher, C. R., & Thurlow, R. (1996). A «landscape» view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. Dans: B. K. Britton & A. C. Graesser (éd.), *Models of understanding text* (p. 165-187). Hillsdale, NJ: Lawrence Erlbaum Associates.
- van den Broek, P., Ridsen, K., & Husebye-Hartmann, E. (1995). The role of readers' standards of coherence in the generation of inferences during reading. Dans: R. F. Lorch, Jr. & E. J. O'Brien (éd.), *Sources of coherence in reading* (p. 353-373). Hillsdale, NJ: Erlbaum.
- van den Broek, P., Virtue, S., Everson, M., Tzeng, Y., & Sung, Y. C. (2002). Comprehension and memory of science texts: Inferential processes and the construction of a mental representation. Dans: J. Otero, J. A. Leon & A. C. Graesser (éd.), *The psychology of science text comprehension* (p. 131-154). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading. Dans: H. van Oostendorp & S. R. Goldman (éd.), *The construction of mental representations during reading* (p. 71-98). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Vosniadou, S., & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science*, 18, 123-183.
- Voss, J., & Silfies, L. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction*, 14, 45-68.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. Dans: F. E. Weinert (éd.), *Leistungsmessung in Schulen* (p. 17-31). Weinheim: Beltz.
- West, L. H. T., & Pines, A. L. (1985). *Cognitive structure and conceptual change*. Orlando, FL: Academic Press.
- Wetzenstein, E. (2004). *Entwicklung eines Anforderungsprofils für Studierende am Beispiel der Psychologie*. Beitrag anlässlich des Bühler-Kolloquium der TU Dresden.
- Willenberg, H. (éd.) (2007). *Kompetenzhandbuch für den Deutschunterricht: auf der empirischen Basis des DESI-Projekts*. Baltmannsweiler: Schneider Hohengehren.
- Wilson, P. T., & Anderson, R. C. (1986). What they don't know will hurt them: The role of prior knowledge in comprehension. Dans: J. Orasanu (éd.), *Reading comprehension: From research to practice* (p. 31-48). Hillsdale, NJ: Lawrence Erlbaum.
- Wonderlic Personnel Test. (1992). *Wonderlic Personnel Test and Scholastic Level Exam: User's Manual*. Libertyville, IL: Wonderlic Personnel Test, Inc.
- Wright, B. D., & Stone, M. H. (1979). *Best Test Design. Rasch Measurement*. University of Chicago: Mesa Press.
- Zheng, A. Y., Lawhorn, J. K., Lumley, T., & Freeman, S. (2008). Application of Bloom's Taxonomy Debunks the «MCAT Myth». *Science*, 319 (5862), 414-415.
- Zimmer, K. (2006). PISA 2003 – Kompetenzmuster von Jungen und Mädchen in den deutschen Ländern. *Unterrichtswissenschaft* 34 (4), 310-329.

Liste des tableaux et figures

Liste des figures

Figure I.1:	Modèle global pour la déduction des critères d'évaluation	33
Figure I.2:	Enquêtes choisies pour EVAMAR II	34
Figure I.3:	Organigramme du projet EVAMAR II	38
Figure II.1:	Liens entre les disciplines d'études et les branches d'enseignement	49
Figure III.1:	Modèles des compétences langagières de Hans Glinz, dit modèle par couches (1987, cité par Sieber, 1994, p. 41)	73
Figure III.2:	Représentation – par EVAMAR – de trois des six domaines de compétence de DESI pour la langue première «Allemand» (DESI-Konsortium, 2006) avec les niveaux de A à E (n'y figurent pas: l'argumentation, le lexique, l'orthographe)	74
Figure III.3:	Le modèle de compétence de PISA pour la compréhension de l'écrit – avec cinq niveaux de compétence. - Présentation par EVAMAR des trois compétences partielles de lecture (d'après Artelt et al., 2004, p. 144-145).	77
Figure III.4:	Présentation par EVAMAR: le modèle de lecture PISA avec intégration des genres de textes envisagés.	78
Figure III.5:	Niveaux communs de référence: échelle globale (extrait niveau C) du CECR (Conseil de l'Europe 2001, p. 25)	81
Figure III.6:	Tableau des spécifications pour l'auto-évaluation de DIALANG – extrait du niveau C (Conseil de l'Europe, 2001, p. 165-168)	82
Figure III.7:	HarmoS: Un modèle pour la description (Sieber, 2006a, p. 18)	85
Figure III.8:	HarmoS – un exemple: savoir écouter (Sieber, 2006a, p. 19)	85
Figure III.10:	Modèle de base en langue première sur la base du CECR (les surfaces grisées correspondent aux domaines analysés par EVAMAR II)	91
Figure III.11:	Grille de compétences en vue des tests en langue première (avec descripteurs de «capacité de faire»)	94
Figure III.12:	Schéma d'un modèle de compétences à plusieurs dimensions	95
Figure III.13:	Représentation en matrice des dimensions du contenu et des aptitudes	95
Figure III.14:	Le modèle de compétences allemand de la KMK	96
Figure III.15:	Aspects et domaines de compétence dans HarmoS (9 ^e année)	98
Figure III.16:	Matrice des aptitudes attendues et des contenus	100
Figure III.17:	Grille de compétences en mathématique pour EVAMAR II	104
Figure III.18:	Comparaison de la structure de trois manuels standard de biologie en langue allemande au niveau gymnasial (degré secondaire II) (titres originaux des chapitres en allemand)	108
Figure III.19:	Comparaison de la structure de deux manuels standard américains et traduits en allemand destinés aux étudiants débutant en biologie (degré tertiaire) (titres originaux des chapitres en anglais et en allemand)	109

Figure III.20: Nom court et contexte des documents traités ci-après	110
Figure III.21: Classification de la biologie en sous-domaines, fondée sur des ouvrages de référence standard de biologie en langues allemande et anglaise, sur les EPA Biologie allemandes et sur la classification allemande de Ratisbonne	114
Figure III.22: Niveaux de culture scientifique (Scientific Literacy) d'après Bybee (1997)	118
Figure III.23: La grille de compétences Connaissances spécifiques en biologie élaborée pour EVAMAR II	121
Figure IV.1: Stratégies d'apprentissage, objectifs de performance et succès dans les études	127
Figure IV.2: Extrait de la liste des unités de sens codées (mots d'origine étrangère en allemand - «Fremdwörter»)	134
Figure IV.3: Item de vocabulaire construit à partir du texte reproduit dans la figure IV.4 et des unités de sens recensées dans la figure IV.2.	135
Figure IV.4: Reproduction du texte original (en allemand) codé par le concepteur du test avec les unités de sens codées, les registres de langue et le choix opérés pour les items du test (Steiner, 2001, p. 16-17)	136
Figure IV.5: Compréhension détaillée de l'écrit - Item construit à partir du texte reproduit dans la figure IV.4 et des unités de sens recensées dans la figure IV.2.	137
Figure IV.6: Extrait d'une note de cours originale de l'EPFZ («Physik, Felder, Geometrie», p. 30)	140
Figure IV.7: Item correspondant à la note de cours de la figure IV.6	141
Figure IV.8: Page tirée d'une note de cours en sciences des médias et de la communication	143
Figure IV.9: Question du test relative à l'extrait de la note de cours reproduit dans la figure IV.8	144
Figure IV.10: Page extraite d'un manuel de biologie (en l'occurrence en allemand)	145
Figure IV.11: Question du test relative à l'extrait de cours reproduit dans la figure IV.10	146
Figure IV.12: Item Characteristic Curve	153
Figure IV.13: Estimation globale des aptitudes personnelles en langue première	156
Figure IV.14: Exercice facile de vocabulaire (exemple)	157
Figure IV.15: Exercice très facile de compréhension détaillée de l'écrit (exemple)	157
Figure IV.17: Exercice difficile de compréhension générale de l'écrit (exemple)	158
Figure IV.18: Exemple d'exercice très difficile de compréhension générale de l'écrit en format ouvert (codage graduel à cinq niveaux)	159
Figure IV.19: Exemple d'exercice difficile de compréhension générale de l'écrit en format ouvert (codage graduel à quatre niveaux)	160
Figure IV.20: Exercice de vocabulaire de difficulté moyenne (exemple)	161
Figure IV.21: Estimation globale des aptitudes personnelles en mathématiques	162
Figure IV.22: Exercice de mathématiques facilement résolu (exemple 1)	163
Figure IV.23: Exercice de mathématiques facilement résolu (exemple 2)	164
Figure IV.24: Exercice de mathématiques facilement résolu (exemple 3)	165
Figure IV.25: Exercice de mathématiques facilement résolu (exemple 4)	165
Figure IV.26: Exercice de mathématiques facilement résolu (exemple 5)	166

Figure IV.27: Exercice de mathématiques ayant posé problème (exemple 1)	166
Figure IV.28: Exercice de mathématiques ayant posé problème (exemple 2)	166
Figure IV.29: Exercice de mathématiques ayant posé problème (exemple 3)	167
Figure IV.30: Exercice de mathématiques avec des composants de difficulté variable (exemple)	168
Figure IV.31: Estimation globale des aptitudes personnelles à la biologie	170
Figure IV.32: Item facile de biologie (exemple 1)	170
Figure IV.33: Item facile de biologie (exemple 2)	171
Figure IV.34: Item facile de biologie (exemple 3)	171
Figure IV.35: Item facile de biologie (exemple 4)	172
Figure IV.36: Item de difficulté moyenne de biologie (exemple 1)	172
Figure IV.37: Item de difficulté moyenne de biologie (exemple 2)	173
Figure IV.38: Item de difficulté moyenne de biologie (exemple 3)	174
Figure IV.39: Item de difficulté moyenne de biologie (exemple 4)	174
Figure IV.40: Item difficile de biologie (exemple 1)	175
Figure IV.41: Item difficile de biologie (exemple 2)	176
Figure IV.42: Item difficile de biologie (exemple 3)	176
Figure IV.43: Item difficile de biologie (exemple 4)	177
Figure IV.44: Estimation globale des aptitudes personnelles d'après les résultats au TCT	180
Figure V.1: Parts des différentes catégories dans le nombre total de points, N = 65 (54 examens de maturité, 9 examens de préparation à la maturité, 2 examens de biologie en tant qu'option complémentaire; D 58, F 4, I 3)	272
Figure V.2: Parts respectives des tâches de mémoire et de transfert (biologie, gymnase)	273
Figure V.3: Comparaison des exigences cognitives des examens de maturité et des examens universitaires en biologie	276
Figure V.4: Parts respectives des tâches de mémoire et de transfert (biologie, université)	276
Figure V.5: Comparaison de la répartition par potentiels d'activation cognitive des tâches des examens de maturité 2007 en Suisse (hachures obliques) et des items sélectionnés pour l'évaluation EVAMAR II en biologie (hachures horizontales)	279
Figure VI.1: Genèse d'un travail de maturité	293
Figure VI.2: Les conditions de réalisation des travaux de maturité dans le schéma de fonctionnement	298
Figure VI.3: Les élèves dans le schéma de fonctionnement	298
Exemple 1: Règlement du 2 février 2001 relatif à la préparation des travaux de maturité dans les gymnases du canton de Bâle-Ville	308
Exemple 2: Ordonnance du 6 juillet 1999 relative aux études gymnasiales du canton des Grisons.	308
Figure VI.4: Séances spécifiques de préparation au travail de maturité dans les gymnases	311
Figure VI.5: Modalités de choix du sujet (N=30)	312
Figure VI.6: Consignes de forme (nombre de mentions)	313

Figure VI.7: Variance du délai de préparation (mois)	314
Figure VI.8: Pondération de la notation finale	316
Figure VI.9: Satisfaction éprouvée à l'égard de l'encadrement fourni par les enseignants (en pour-cent)	324
Figure VI.10: Fréquence des contacts avec la personne suivant le travail de maturité (en pour-cent)	325
Figure VI.11: Fréquence des entretiens individuels de l'élève avec la personne suivant son travail de maturité (en pour-cent)	325
Figure VI.12: Satisfaction des élèves à l'égard de la fréquence des entretiens (en pour-cent)	326
Figure VI.13: Appréciation de la compétence de l'enseignant sur le fond (en pour-cent)	329
Figure VI.14: Appréciation portée par les élèves sur l'utilité des séances et cours de préparation au travail de maturité (en pour-cent)	331
Figure VI.15: Cadre institutionnel et individuel de préparation du travail de maturité	332
Figure VI.16: Appréciation du volume de travail de préparation et de l'utilité (en pour-cent)	334
Figure VI.17: Appréciation de l'utilité du travail de maturité dans l'itinéraire qualifiant et le développement personnel	335
Figure VII.1: Enquêtes retenues pour EVAMAR II	367
Suite figure VII.1: Enquêtes retenues pour EVAMAR II	368
Figure VII.2: Relation entre disciplines de maturité et branches d'études	369

Liste des tableaux

Tableau II.1: Disciplines d'études retenues dans la base de données du module A	41
Tableau II.2: Grille prescrite pour saisir la dimension cognitive et les disciplines scolaires	43
Tableau II.3: Catégories de connaissance des unités de sens (catégories principales)	44
Tableau II.4: Catégories d'aptitudes cognitives pour les tâches et les examens	45
Tableau II.5: Importance des branches gymnasiales pour le propre enseignement	52
Tableau II.6: Importance des branches gymnasiales en fonction des disciplines d'études (classement par ordre d'importance entre parenthèses)	54
Tableau II.7: Connaissances et capacités spécifiques requises: mentions expresses des attentes par discipline	56
Tableau II.8: Connaissances et capacités spécifiques requises: mentions expresses des lacunes par discipline	59
Tableau II.9: Importance des compétences interdisciplinaires pour le propre enseignement	62
Tableau II.10: Carences dans les connaissances et capacités interdisciplinaires jugées importantes	63
Tableau II.11: Connaissances et capacités supplémentaires souhaitables au début des études	64
Tableau IV.1: Composantes importantes d'une maturité universitaire élevée	128
Tableau IV.2: Les sept cahiers de tests en biologie	130
Tableau IV.3: Population cible	151
Tableau IV.4: Echantillon	151

Tableau IV.5:	Niveau de participation des classes et des élèves selon la strate géographique	151
Tableau IV.6:	Résultats du test de langue première	154
Tableau IV.7:	Résultats du test de mathématiques	161
Tableau IV.8:	Résultats du test de biologie	169
Tableau IV.9:	Sélection de notes de maturité	178
Tableau IV.10:	Résultat du test des compétences transversales (TCT)	179
Tableau IV.11:	Valeurs moyennes sur les échelles des compétences transversales	181
Tableau IV.12:	Valeurs moyennes des questions sur la persévérance (1 = très rarement; 5 = très souvent)	182
Tableau IV.13:	Intérêt pour la matière enseignée	182
Tableau IV.14:	Comparaison des résultats au test de langue première selon la strate géographique (total)	186
Tableau IV.15:	Comparaison des résultats au test de langue première selon la strate géographique (compréhension générale de l'écrit)	186
Tableau IV.16:	Comparaison des résultats au test de langue première selon la strate géographique (compréhension détaillée de l'écrit)	186
Tableau IV.17:	Comparaison des résultats au test de langue première selon la strate géographique (fonctionnement de la langue/vocabulaire)	187
Tableau IV.18:	Comparaison des résultats au test de langue première selon la strate géographique (fonctionnement de la langue/compétence grammaticale et orthographe)	187
Tableau IV.19:	Comparaison des résultats au test de mathématiques selon la strate géographique (total)	188
Tableau IV.20:	Comparaison des résultats au test de mathématiques selon la strate géographique (calcul analytique)	188
Tableau IV.21:	Comparaison des résultats au test de mathématiques selon la strate géographique (géométrie)	188
Tableau IV.22:	Comparaison des résultats au test de mathématiques selon la strate géographique (stochastique)	189
Tableau IV.23:	Comparaison des résultats au test de mathématiques selon la strate géographique (divers)	189
Tableau IV.24:	Comparaison des résultats au test de biologie selon la strate géographique (total)	190
Tableau IV.25:	Comparaison des résultats au test de biologie selon la strate géographique (évolution)	190
Tableau IV.26:	Comparaison des résultats au test de biologie selon la strate géographique (structure et fonction)	190
Tableau IV.27:	Comparaison des résultats au test de biologie selon la strate géographique (génétique)	191
Tableau IV.28:	Comparaison des résultats au test de biologie selon la strate géographique (traitement de l'information)	191
Tableau IV.29:	Comparaison des résultats au test de biologie selon la strate géographique (écologie)	191
Tableau IV.30:	Comparaison des résultats au test de biologie selon la strate géographique (métabolisme)	191
Tableau IV.31:	Comparaison des résultats au test de biologie des élèves qui avaient déjà terminé les cours dans cette discipline selon la strate géographique (total)	192

Tableau IV.32: Comparaison des résultats au test de biologie des élèves qui n'avaient pas encore terminé les cours dans cette discipline selon la strate géographique (total)	192
Tableau IV.33: Comparaison des notes de maturité en langue première selon la strate géographique (note finale)	193
Tableau IV.34: Comparaison des notes de maturité en mathématiques selon la strate géographique (note finale)	193
Tableau IV.35: Comparaison des notes de maturité en biologie selon la strate géographique (notation continue)	193
Tableau IV.36: Comparaison des notes de maturité en sciences expérimentales selon la strate géographique (note finale)	194
Tableau IV.37: Comparaison des résultats du test des compétences transversales selon la strate géographique (TCT, quantités et formes)	194
Tableau IV.38: Comparaison des résultats du test des compétences transversales selon la strate géographique (TCT, diagrammes et tableaux)	194
Tableau IV.39: Comparaison des résultats du test des compétences transversales selon la strate géographique (TCT, compréhension de textes)	195
Tableau IV.40: Comparaison des résultats du test des compétences transversales selon la strate géographique (TCT, notions élémentaires de sciences expérimentales)	195
Tableau IV.41: Comparaison des résultats de l'auto-évaluation des compétences transversales selon la strate géographique (planification du temps de travail: 1 = mauvaise, 5 = très bonne)	195
Tableau IV.42: Comparaison des résultats de l'auto-évaluation des compétences transversales selon la strate géographique (efficacité de la préparation aux examens: 1 = préparation inefficace, 5 = préparation très efficace)	196
Tableau IV.43: Comparaison des résultats de l'auto-évaluation des compétences transversales selon la strate géographique (gestion du temps: 1 = très mauvaise, 5 = très bonne)	196
Tableau IV.44: Comparaison des résultats au test de langue première selon l'option spécifique choisie (total)	197
Tableau IV.45: Comparaison des résultats au test de langue première selon l'option spécifique choisie (compréhension générale de l'écrit)	197
Tableau IV.46: Comparaison des résultats au test de langue première selon l'option spécifique choisie (compréhension détaillée de l'écrit)	197
Tableau IV.47: Comparaison des résultats au test de langue première selon l'option spécifique choisie (fonctionnement de la langue/vocabulaire)	198
Tableau IV.48: Comparaison des résultats au test de langue première selon l'option spécifique choisie (fonctionnement de la langue/compétence grammaticale et orthographe)	198
Tableau IV.49: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (total)	199
Tableau IV.50: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (calcul analytique)	199
Tableau IV.51: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (géométrie)	199

Tableau IV.52: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (stochastique)	200
Tableau IV.53: Comparaison des résultats au test de mathématiques selon l'option spécifique choisie (divers)	200
Tableau IV.54: Comparaison des résultats au test de biologie selon l'option spécifique choisie (total)	201
Tableau IV.55: Comparaison des résultats au test de biologie selon l'option spécifique choisie (évolution)	201
Tableau IV.56: Comparaison des résultats au test de biologie selon l'option spécifique choisie (structure et fonction)	201
Tableau IV.57: Comparaison des résultats au test de biologie selon l'option spécifique choisie (génétique)	202
Tableau IV.58: Comparaison des résultats au test de biologie selon l'option spécifique choisie (traitement de l'information)	202
Tableau IV.59: Comparaison des résultats au test de biologie selon l'option spécifique choisie (écologie)	202
Tableau IV.60: Comparaison des résultats au test de biologie selon l'option spécifique choisie (métabolisme)	203
Tableau IV.61: Comparaison des notes de maturité selon l'option spécifique choisie	203
Tableau IV.62: Comparaison des notes de maturité en mathématiques selon l'option spécifique choisie	204
Tableau IV.63: Comparaison des notations continues comptant pour la maturité en biologie selon l'option spécifique choisie	204
Tableau IV.64: Comparaison des notes de maturité en sciences expérimentales selon l'option spécifique choisie	205
Tableau IV.65: Comparaison de la proportion de notes insuffisantes aux examens de maturité selon l'option spécifique choisie	205
Tableau IV.66: Comparaison des résultats du test des compétences transversales selon l'option spécifique choisie (TCT, quantités et formes)	206
Tableau IV.67: Comparaison des résultats du test des compétences transversales selon l'option spécifique choisie (TCT, diagrammes et tableaux)	206
Tableau IV.68: Comparaison des résultats du test des compétences transversales selon l'option spécifique choisie (TCT, compréhension des textes)	207
Tableau IV.69: Comparaison des résultats du test des compétences transversales selon l'option spécifique choisie (TCT, notions élémentaires de sciences expérimentales)	207
Tableau IV.70: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (planification du temps de travail: 1 = mauvaise, 5 = très bonne)	208
Tableau IV.71: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (efficacité de la préparation aux examens: 1 = inefficace, 5 = très efficace)	208
Tableau IV.72: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (gestion du temps: 1 = très mauvaise, 5 = très bonne)	209
Tableau IV.73: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (capacité d'analyse et de synthèse: 1 = très faible, 5 = très développée)	209

Tableau IV.74: Comparaison des résultats de l'auto-évaluation des compétences transversales selon l'option spécifique choisie (capacité d'adopter d'autres points de vue en cas de conflit: 1 = très bonne, 4 = très mauvaise)	209
Tableau IV.75: Ecart en langue première selon le sexe	210
Tableau IV.76: Ecart en mathématiques selon le sexe	211
Tableau IV.77: Ecart en biologie et en sciences naturelles selon le sexe	212
Tableau IV.78: Ecart dans les résultats du TCT selon le sexe	213
Tableau IV.79: Domaines du TCT où les écarts sont significatifs entre les sexes (amplitude de l'effet > 0,2)	213
Tableau IV.80: Ecart en langue première selon le taux de maturité dans les groupes de cantons	214
Tableau IV.81: Ecart en mathématiques selon le taux de maturité dans les groupes de cantons	214
Tableau IV.82: Ecart en sciences expérimentales selon le taux de maturité dans les groupes de cantons	215
Tableau IV.83: Ecart dans le TCT selon le taux de maturité dans les groupes de cantons	215
Tableau IV.84: Ecart en langue première selon la durée de la voie gymnasiale	216
Tableau IV.85: Ecart en mathématiques selon la durée de la voie gymnasiale	217
Tableau IV.86: Ecart en sciences expérimentales selon la durée de la voie gymnasiale	217
Tableau IV.87: Ecart dans le TCT selon la durée de la voie gymnasiale	218
Tableau IV.88: Lien entre l'application des élèves lors des tests et leurs résultats de mathématiques	219
Tableau IV.89: Lien entre l'application des élèves lors des tests et leur notations continues en mathématiques	219
Tableau IV.90: Lien entre l'application des élèves lors des tests et leurs notes de maturité à l'examen écrit en mathématiques	219
Tableau IV.91: Lien entre l'application des élèves lors des tests et leurs notes finales en mathématiques	220
Tableau IV.92: Comparaison des notes obtenues à l'examen oral et écrit	220
Tableau IV.93: Corrélations entre les tests	221
Tableau IV.94: Corrélations entre le test et les notes de maturité en langue première	222
Tableau IV.95: Corrélations entre le test et les notes de maturité en mathématiques	222
Tableau IV.96: Corrélations entre le test et les notes de maturité en biologie	222
Tableau IV.97: Corrélations entre le TCT et les notes de maturité	223
Tableau IV.98: Corrélations des notes de maturité entre elles	223
Tableau V.1: Tableau des catégories et sous-catégories des processus cognitifs selon Anderson et al. (2001) et des définitions adaptées dans le cadre d'EVAMAR II (appliquée pour l'instant seulement à la biologie).	229
Tableau V.2: Les formes de composition à l'examen de maturité dans les gymnases suisses alémaniques	234
Tableau V.3: Les formes de composition à l'examen de maturité en Suisse italienne	235
Tableau V.4: Les formes de composition en Suisse romande	235

Tableau V.5:	Les types de textes soumis aux élèves dans les épreuves de maturité de Suisse alémanique	237
Tableau V.6:	Les types de textes soumis aux élèves dans les épreuves de maturité de Suisse romande	238
Tableau V.7:	Les types de textes soumis aux élèves dans les épreuves de maturité de Suisse italienne (Tessin)	239
Tableau V.8:	Indications des siècles auxquels appartiennent les textes soumis aux élèves (épreuves de maturité en Suisse, trois régions linguistiques)	240
Tableau V.9:	Auteurs cités dans les épreuves de maturité de Suisse alémanique (extrait)	240
Tableau V.10:	Origine/nationalité des auteurs cités (épreuves de maturité d'allemand de Suisse alémanique)	241
Tableau V.11:	Origine/nationalité des auteurs cités (épreuves de maturité de français de Suisse romande)	242
Tableau V.12:	Genre (sexe) des auteurs cités (données des épreuves de maturité de toute la Suisse)	242
Tableau V.13:	Longueur des textes de départ soumis (épreuves écrites de maturité, les trois régions linguistiques)	243
Tableau V.14:	Nombre de textes de départ soumis (épreuves écrites de maturité, les trois régions linguistiques)	244
Tableau V.15:	Nombre de sujets de composition proposés à choix (épreuves de maturité de Suisse alémanique et de Suisse romande)	244
Tableau V.16:	Explications complémentaires de l'enseignant (épreuves de maturité de Suisse alémanique)	245
Tableau V.17:	Moyens auxiliaires autorisés (épreuves de maturité de Suisse romande, cumul de mentions possible)	246
Tableau V.18:	Analyse des exigences cognitives (épreuves de maturité de Suisse alémanique, catégories principales)	247
Tableau V.19:	Analyse des exigences cognitives (catégories principales et sous-catégories, pondérées; double désignation; épreuves de maturité de Suisse alémanique)	248
Tableau V.20:	Analyse des exigences cognitives (examens de maturité au Tessin)	249
Tableau V.21:	Analyse des exigences cognitives (examens de maturité en Suisse romande)	249
Tableau V.22:	Modèle des potentiels d'activités cognitives pour les mathématiques	253
Tableau V.23:	Répartition (en pour-cent) des items selon les divers chapitres des mathématiques	255
Tableau V.24:	Répartition (en pour-cent) des problèmes regroupés selon les principaux chapitres	256
Tableau V.25:	Exigences cognitives des items ou des problèmes	257
Tableau V.26:	Représentation quantitative des exigences cognitives potentielles	262
Tableau V.27:	Comparaison des exigences des professeurs avec les examens de maturité	263
Tableau V.28:	Domaines thématiques pour l'analyse des épreuves de maturité en biologie	267
Tableau V.29:	Poids moyen du domaine thématique t dans la note ou le nombre total de points disponibles	268

Tableau V.30: Part du nombre total de points pouvant être obtenus par la résolution de questions d'une certaine catégorie dans un examen de maturité moyen en biologie.	270
Tableau V.31: Répartition du nombre total de points dans les examens universitaires en biologie par catégories de potentiel d'activation cognitive	274
Tableau V.32: Répartition des items sélectionnés pour le test de biologie dans EVAMAR II par catégories de potentiel d'activation cognitive	278
Tableau V.33: Déficits constatés en biologie et savoir et capacités supplémentaires souhaités chez les nouveaux étudiants (en italique: écologie et systématique/évolution)	280
Tableaux VI.x: voir chapitre VI	

Liste des abréviations

Paramètres et concepts du modèle de Rasch

θ	aptitudes personnelles
δ	difficulté de l'item
ICC	Item Characteristic Curve
IRC	Item Response Curve
p	probabilité de résolution

Paramètres des comparaisons multiples de moyennes

d	amplitude de l'effet
d max	amplitude de l'effet lorsque l'écart entre les moyennes est le plus grand (amplitude maximale)

Disciplines et domaines d'enseignement gymnasial, sous-projet A1

All	allemand
Ang	anglais
Av	arts visuels
B	biologie
C	chimie
Div	divers
ED	économie et droit
F	français
G	géographie
Gr	grec ancien
H	histoire
I	italien
Inf	informatique et nouveaux médias
L	latin
M	mathématiques
Mu	musique
P	physique
PPP	philosophie, pédagogie, psychologie, éthique et enseignement religieux
Ts	Travaux scientifiques

Domaines de compétence en langue première

C-E_DETAIL	compréhension détaillée de l'écrit (information et argumentation)
C-E_GEN	compréhension générale de l'écrit (lire pour s'informer)

GRAM	fonctionnement de la langue/compétence grammaticale et orthographe
TOT_LANGUE	résultat global en langue première (hors grammaire)
VOC	fonctionnement de la langue/vocabulaire

Domaines de compétence en mathématique

AF	calcul analytique/fonctions
AI	calcul analytique/calcul infinitésimal
AR	calcul analytique/suites et séries
GG	géométrie/géométrie élémentaire et analytique
GT	géométrie/trigonométrie
GV	géométrie/calcul vectoriel et géométrie vectorielle
SK	stochastique/analyse combinatoire
SS	stochastique/statistique
SW	stochastique/calcul des probabilités
XA	divers/algèbre élémentaire
XL	divers/logique
XM	divers/théorie des ensembles
ANALYSE	calcul analytique (fonctions, calcul infinitésimal, suites et séries)
GEOMETRIE	géométrie (trigonométrie, géométrie analytique, calcul vectoriel)
DIVERS_M	divers (théorie des ensembles/mathématiques discrètes, logique, algèbre élémentaire)
STOCHASTIQUE	stochastique (calcul des probabilités, analyse combinatoire, statistique)
TOT_MATH	résultat global en mathématiques

Domaines de compétence en biologie

B_ECO	écologie (approche holistique, écologie et développement durable)
B_EVOL	évolution (processus de développement, évolution et projection)
B_STRU	structure et fonction (construction et fonction des cellules, tissus et organes; différenciations fonctionnelles)
B_GEN	génétique (bases de génétique moléculaire, application des connaissances et méthodes de la biologie moderne)
B_INFO	traitement de l'information (communication entre cellules, traitement neuronal de l'information, perception)
B_MET	métabolisme (transformations moléculaires et énergétiques: nécessité et voies de transformation énergétique, anabolisme et catabolisme relatifs aux structures cellulaires et de l'organisme; régulation moléculaire du métabolisme: enzymologie)
TOT_BIOL	résultat global en biologie

Domaines de compétence du test des compétences transversales

T_DIAGR	diagrammes et tableaux
T_QUANT	ensembles et formes
T_SCIEN	notions élémentaires de sciences expérimentales
T_TEXT	compréhension de textes

Paramètres du questionnaire complémentaire

Pruefvo	préparation à l'examen
Zeiteff	efficience dans la gestion du temps à disposition
Zeitpla	planification du temps à disposition

Notes en langue première

E_ErfNote	langue première: notation continue
E_MatNote	langue première: note finale
E_müPr	langue première: note à l'examen oral de maturité
E_schrPr	langue première: note à l'examen écrit de maturité

Notes en mathématiques

Ma_ErfNote	mathématiques: notation continue
Ma_MatNote	mathématiques: note finale
Ma_müPr	mathématiques: note à l'examen oral de maturité
Ma_schrPr	mathématiques: note à l'examen écrit de maturité
MEF_schrPr	mathématiques, option complémentaire (OC): note à l'examen écrit de maturité
MEF_müPr	mathématiques, option complémentaire (OC): note à l'examen oral de maturité
MPhys_MatPr	physique et applications des mathématiques, option spécifique (OS): note finale
MphyS_müPr	physique et applications des mathématiques, option spécifique (OS): note à l'examen oral de maturité
MphyS_schrPr	physique et applications des mathématiques, option spécifique (OS): note à l'examen écrit de maturité
Ph_ErfNote	physique: notation continue

Notes en biologie

Bi_ChS_schrPr	biologie et chimie, option spécifique (OS): note à l'examen écrit de maturité
Bi_ErfNote	biologie: notation continue
Bi_müPr	biologie: note à l'examen oral de maturité
Bi_schrPr	biologie: note à l'examen écrit de maturité
BiChS_müPr	biologie et chimie, option spécifique (OS): note à l'examen oral de maturité

BiEF_müPr	biologie, option complémentaire (OC): note à l'examen oral de maturité
BiEF_schrPr	biologie, option complémentaire (OC): note à l'examen écrit de maturité
Ch_ErfNote	chimie: notation continue
NW_MatNote	sciences expérimentales: note finale

Cantons suisses

AG	Argovie
AI	Appenzell Rhodes-Intérieures
AR	Appenzell Rhodes-Extérieures
BE _d	Berne, partie germanophone
BE _f	Berne, partie francophone
BS	Bâle-Ville
FR	Fribourg
GL	Glaris
GR	Grisons
JU	Jura
LU	Lucerne
NE	Neuchâtel
NW	Nidwald
OW	Obwald
SG	Saint-Gall
SH	Schaffhouse
SO	Soleure
SZ	Schwyz
TG	Thurgovie
UR	Uri
VD	Vaud
VS	Valais
VS _d	Valais, partie germanophone
VS _f	Valais, partie francophone
ZG	Zoug

Abréviations générales

ADN	acide désoxyribonucléique
ADN	acide désoxyribonucléique
ALTE	Association des centres d'évaluation en langue en Europe (<i>Association of Language Teachers in Europe</i>)
AMS	test d'aptitudes pour les études de médecine en Suisse (aux universités de Bâle, Berne, Fribourg et Zurich)
anc.	anciennement

ATP	adénosine triphosphate
BCH	option spécifique Biologie et chimie
bik	projet <i>Biologie im Kontext</i> («Biologie en contexte»), BMBF
BMBF	Ministère fédéral allemand de la formation et de la recherche (<i>Bundesministerium für Bildung und Forschung</i>)
c.-à-d.	c'est-à-dire
CDGS	Conférence des directeurs des gymnases suisses
CDIP	Conférence suisse des directeurs cantonaux de l'instruction publique
CECR	Cadre européen commun de référence pour les langues
cf.	<i>conferre</i> («voir aussi...»)
chap.	chapitre
CRUS	Conférence des recteurs des universités suisses
CTD	Centre pour le développement de tests et le diagnostic, Université de Fribourg
DaF	allemand langue étrangère (<i>Deutsch als Fremdsprache</i>)
DESI	<i>Deutsch-Englisch-Schülerleistungen-International</i> (programme fédéral de recherche didactique sur l'enseignement des langues en Allemagne)
éd.	éditeur(s)
EPA	Exigences communes pour l'examen de baccalauréat allemand en biologie (<i>Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie</i>)
EPFL	Ecole polytechnique fédérale de Lausanne
EPFZ	Ecole polytechnique fédérale de Zurich
et al.	<i>et alii</i> («et les autres»: et coauteurs)
étud.	étudiants
EVAMAR I	Evaluation de la réforme de la maturité 1995, phase I (2004)
EVAMAR II	Evaluation de la réforme de la maturité 1995, phase II (2008)
FHNW	Haute école spécialisée (HES) du Nord-Ouest de la Suisse (<i>Fachhochschule Nordwestschweiz</i>)
GRE	Biology Graduate Record Examination (Etats-Unis)
HarmoS	accord intercantonal sur l'harmonisation de la scolarité obligatoire
HEP	haute école pédagogique
HES	haute école spécialisée
HHS	heures hebdomadaires (de cours) par semestre
HRM	<i>Human Resources Management</i> (gestion du personnel)
HSG	Université de Saint-Gall
IBE	Institut für Bildungsevaluation, Université de Zurich
IGB	Institut für Gymnasial- und Berufspädagogik, Université de Zurich
IPN	<i>Leibniz-Institut für die Pädagogik der Naturwissenschaften</i> , Kiel
IRDP	Institut de recherche et de documentation pédagogique, Neuchâtel
KMK	Conférence permanente des ministres de l'éducation et des affaires culturelles des länder allemands (<i>Kultusministerkonferenz</i>)
maths	mathématiques

MCAT	Medical College Admission Test (Etats-Unis)
min.	minute(s)
MNU	Société allemande pour la promotion de l'enseignement des mathématiques et des sciences naturelles (<i>Deutscher Verein zur Förderung des mathematischen und naturwissenschaftlichen Unterrichts e.V. [=eingetragener Verein]</i>)
MSWNRW	Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen
NCTM	National Council of Teachers of Mathematics (Etats-Unis)
n. i.	non indiqué
NRW	Rhénanie-du-Nord-Westphalie (<i>Nordrhein-Westfalen</i>)
OC	option complémentaire
OCDE	Organisation de coopération et de développement économiques
OLAT	Online Learning And Training (plateforme de formation en ligne gérée par l'Université de Zurich)
ORM 68	ordonnance sur la reconnaissance de certificats de maturité de 1968
OS	option spécifique
p.	page(s)
PAM	option spécifique Physique et applications des mathématiques
par ex.	par exemple
PEC	plan d'études cadre (de la CDIP)
PH	Haute école pédagogique, HEP (<i>Pädagogische Hochschule</i>)
PISA	Programme international pour le suivi des acquis des élèves
PPP	option spécifique Philosophie/pédagogie/psychologie
prés.	présentation
QCM	question(s) à choix multiple
RRM 95	règlement sur la reconnaissance des certificats de maturité gymnasiale de 1995
RVK	classification allemande de Ratisbonne (<i>Regensburger Verbundklassifikation</i>)
s	et suivante (après un numéro de page)
SD	écart type (<i>standard deviation</i>)
SER	Secrétariat d'Etat à l'éducation et à la recherche
SI	degré secondaire I
SII	degré secondaire II
SP	sous-projet (dans EVAMAR II)
ss	et suivantes (après un numéro de page)
SSPES	Société suisse des professeurs de l'enseignement secondaire
TCT	test des compétences transversales
TIMSS	Etudes des tendances internationales en mathématiques et en sciences (<i>Trends in International Mathematics and Science Study</i> ; anc. <i>Third International Mathematics and Science Study</i>)
TMS	test d'aptitudes pour les études de médecine en Allemagne (<i>Test für medizinische Studiengänge</i>)

TPE	tableau périodique des éléments
USR	Ufficio studi e ricerca, Bellinzona
vs	versus (par opposition à)
VSEPR	répulsion des paires électroniques de la couche de valence (<i>Valence shell electron pair repulsion</i>)



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Département fédéral de l'intérieur DFI
Secrétariat d'Etat à l'éducation et à la recherche SER
Education

Hallwylstrasse 4
CH - 3003 Berne
T +41 31 322 96 57
F +41 31 322 78 54
info@sbf.admin.ch
www.sbf.admin.ch